GENETICS OF ALUMINUM TOLERANCE IN MAIZE

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Allison M Krill

August 2013

ABSTRACT

Aluminum (Al) toxicity is a major agronomic problem on acid soils worldwide, due to increased solubility of Al^{3+} at a low pH, and inhibition of plant growth. While many crops are tolerant or adaptable to acid soils, important grain crops such as maize are less tolerant and can result in low yields. Within the genetic diversity of maize we see a wide range of tolerance levels to acid soils and aluminum toxicity. This diversity can be used to improve maize cultivars and better understand the genetics and physiology underlying tolerance mechanisms. This thesis provides an overview of Al toxicity and tolerance studies in maize and other crops, as well as novel phenotype-genotype association results in maize. In order to improve crop yields and fertility of cropland a combination of improved genetics, sustainable management practices, and amending of acid soils should be used.

BIOGRAPHICAL SKETCH

Education

M.S. Plant Breeding and Genetics, August 2013
Cornell University, Dept of Plant Breeding and Genetics (Ithaca, NY)

B.S. Environmental Science/Biology, Minor in Chemistry, May 2003
B.A. Music/piano, May 2003
Kutztown University of Pennsylvania (Kutztown, PA)

Work experience

Graduate Research Assistant, March 2007- May 2009
Dr. Edward S Buckler, USDA-ARS, Cornell University, Institute for Genomic Diversity (Ithaca, NY) *Genetics of aluminum tolerance in maize*

Research Technician, May 2004 - March 2007
Dr. Edward S Buckler, USDA-ARS, Cornell University, Institute for Genomic Diversity, (Ithaca, NY) *Maize genetics, molecular biology, laboratory management, greenhouse and physiology research for Aluminum tolerance.*

Research Intern, March 2003 – May 2004
The Rodale Institute (Kutztown, PA) *Applied research pertaining to sustainable agriculture systems, safety regulation*

Research Assistant, Sept 2002 – May 2003
Dr. Kestas Bendiskas, Dept of Chemistry, Kutztown University of PA (Kutztown, PA)
Development and publication of a biochemistry lab for undergraduate education

Research Intern, May 2002- Sept 2002
State University of NY, Plattsburg (Plattsburg, NY) *Forest Ecology research and GIS database development*

Published work

Krill AM, Kirst M, Kochian LV, Buckler ES, Hoekenga OA (2010) Association and Linkage Analysis of Aluminum Tolerance Genes in Maize. *PLoS ONE* 5(4): e9958.

Bendinskas K, DiJiacomo C, **Krill A**, Vitz E (2005) Kinetics of Alcohol Dehydrogenase-Catalyzed Oxidation of Ethanol Followed by Visible Spectroscopy. *Journal of Chemical Education 82* (7), 1068

Krill A, Newell C, O'Neil M, Phommala O, Adams K. (2004) Plant Community Succession Following Disturbances in a Pine Barren and Adjacent Hardwood Forest. *Scientia discipulorum* 2004:1-16

Presented research

50th annual maize genetics conference, 2008 (Washington DC), Poster
Using nested association mapping to study aluminum tolerance in maize

49th annual maize genetics conference, 2007 (St. Charles, IL), Poster
Analysis of Significant Polymorphisms within Candidate Aluminum Tolerance Genes

48th annual maize genetics conference, 2006 (Pacific Grove, CA), Poster
Association Analysis of Candidate Genes for Aluminum Tolerance in Maize

Grants/awards

Toward Sustainability Foundation Grant, Cornell University, 2008
Effects of maize diversity in organic/non-organic cropping systems at the Rodale institute

Travel Grant, International Agriculture/developing nations, Cornell University, 2007

Scholarship, Summer Institute in Statistical Genetics, University of Washington, 2007

NSF Undergraduate Research Fellowship, SUNY Plattsburg, 2002

This Thesis is dedicated to my wonderful husband and children, Patrick Brown, Cyrus Brown and Alexander Brown

ACKNOWLEDGMENTS

Edward S Buckler

Owen Hoekenga

Leon Kochian

TABLE OF CONTENTS

Abstract

Chapter 1. Aluminum toxicity

Acid soils Page 1
Effects and mechanisms of Al^{3+} toxicity Page 4
Al^{3+} tolerance/resistance mechanisms in plants Page 5
Genetics of Al^{3+} tolerance in plants Page 8
QTL studies of Al^{3+} tolerance in maize Page 12
Expression/physiology studies of Al^{3+} tolerance in maize Page 14
Al^{3+} tolerance genes in maize Page 16
References Page 18

Chapter 2 . Association and Linkage Analysis of Aluminum Tolerance Genes in Maize

Introduction Page 22
Results
- Phenotypic data Page 25
- Genotypic data Page 26
- Association Mapping Page 27
- Linkage Mapping Page 28
- Al tolerance gene: *ZmASL* Page 31
- Al tolerance gene: *SAHH* Page 32
- Al tolerance gene: *ME* Page 33
- Al tolerance gene: *ALMT2* Page 35

Discussion Page 36
Materials and Methods Page 38
References Page 43

Chapter 3. Nested Association Mapping for Aluminum Tolerance

Introduction Page 48
Materials and Methods Page 49
Results Page 50
Discussion Page 54
Conclusion Page 55
Future of Acid soils Page 58
References Page 60

LIST OF FIGURES

Chapter 1

Figure 1. World Soil pH (FAO-UNESCO) Page 1

Figure 2. Effect of pH on bioavailability of minerals Page 2

Figure 3. Effect of pH on aluminum species Page 3

Figure 4. Page 6
(A) Internal and (B) Exclusion mechanisms for Al^{3+} tolerance in plants. Plasma membrane (PM). **A. Internal tolerance.** (1) Binding of Al^{3+} by ligands to form harmless complexes. (2) Storage of these complexes in vacuoles or other organelles. (3) Rapid repair to the cell wall and membranes. (4) scavenging of reactive oxygen species (ROS). **B. Exclusion mechanisms.** (1) Lowering Al^{3+} concentration in cytosol, by active export or vesicular transport. (2) Transport processes across PM that raise apoplastic pH and reduce molar fraction of Al^{3+} (3) Efflux of ligands that bind with Al^{3+} to form harmless complexes (4) Lowered charge density on membrane.

Chapter 2

Figure 1. Page 25
Distribution of Net Root Growth (NRG) in the Maize Association Panel. Bins reflect grouping the inbred lines by 5 mm 2 d^{-1} increments of root growth under both Al stress (circles) and control (squares) treatments. Reported values are Least Squares Means calculated from five replicate experiments for Al stress treatment or three replicate experiments for control treatment.

Figure 2. Page 30
Linkage mapping validation of candidate Al tolerance genes. Six candidate Al tolerance genes were evaluated using three F2 linkage populations: B73xCML247, B73xCML333, and B73xNC350. Mean Net Root Growth values for each allelic state are reported, abbreviated as B73 for the B73 homozygous class, Het for the heterozygous class, and the numerical portion of the non-B73 parent name for the other homozygous class; error bars reflect standard error. Student's t-test was used to evaluate differences between allelic classes within each F_2; differences significant at $p<0.05$ are indicated with letter codes.

Figure 3. Page 32
Characterization of *ZmASL*. (a) Predicted gene structure for the *ZmASL* locus is shown, with exons as black boxes, introns as thin lines, and UTRs as open boxes. The approximate location of 12 amino acid substitutions or additions that occur at greater than 10% frequency among alleles

are shown with vertical lines above the exons, based on complete *ZmASL* sequencing performed in the 27 DL subset. (b) A focus region of *ZmASL* was sequenced in the association panel. The polymorphisms that were identified as significant by the association analyses are shown – SNPs as vertical lines, indels as triangles – and are referred to by number. The conserved MATE domain is highlighted in exons 3 and 4. (c) Linkage disequilibrium plot for the eleven significant polymorphisms. High linkage disequilibrium exists between nine of the eleven associated polymorphisms. (d) An estimate for the transmembrane structure of *ZmASL*, where open circles indicate the approximate locations for the 12 amino acid substitution/insertions detected within the gene.

Figure 4. Page 33
Characterization of *SAHH*.
(a) Predicted gene structure for the *SAHH* locus is shown, with exons as black boxes, introns as thin lines, and UTRs as open boxes. A single amino acid substitution was detected from complete gene sequencing in the 27 DL subset and is indicated by the vertical line in the first exon. (b) A focus region of *SAHH* was sequenced in the association panel. Six SNPs were detected in the association panel and are referred to by number. Polymorphisms 1, 2 and 5 were identified as significantly associated with aluminum tolerance differences and are shown in black; non-significant sites are shown in gray. Site #5 corresponds to the triallelic SNP that causes the single amino acid substitution detected. (c) Linkage disequilibrium plot for all polymorphisms detected in the focus region at *SAHH*. High linkage disequilibrium exists between sites 1 and 2, while relatively low linkage disequilibrium exists through the rest of the gene.

Figure 5. Page 34
Characterization of ME.
(a) Predicted gene structure for the *ME* locus is shown, with exons as black boxes, introns as thin lines, and UTRs as open boxes. Two indels and an SNP were detected in sequencing the complete gene in 27 DL subset. (b) A focus region of *ME* was sequenced in the association panel. Four indels and eight SNPs were detected; three SNPs were significantly associated with aluminum tolerance and are shown in black. (c) Linkage disequilibrium plot for all polymorphisms detected within the focus region at *ME*. No linkage disequilibrium exists within the 5' end of the focus region, while moderate linkage disequilibrium exists among several of the 3' end sites.

Figure 6. Page 35
Characterization of *ZmALMT2*. (a) Predicted gene structure for the *ZmALMT2* locus is shown, with exons as black boxes, introns as thin lines, and UTRs as open boxes. For ease of presentation, polymorphisms detected in the 27 DL complete gene sequencing are not shown. (b) A focus region of *ZmALMT2* was sequenced in the association panel. Four indels and thirteen SNPs were detected; two SNPs were significantly associated with aluminum tolerance and are shown in black (#2, #12). (c) Linkage disequilibrium plot for all polymorphisms detected within the focus region at *ZmALMT2*. Site 2 is associated with aluminum tolerance differences but is in linkage

equilibrium with all other sites, while site 12 has moderate linkage disequilibrium with many sites within the focus region.

Chapter 3

Figure 1. Page 51
Histograms showing the distribution of traits, NRG, RRG, and BLUP_NRG, in the entire NAM population.

Figure 2. Page 51
Phenotypic distributions by population for NRG, BLUP-NRG and RRG and number of lines used in each.

Figure 3. Page 53
Significant markers ($p<0.01$) for NRG, RRG, and BLUP estimates (shown as –log p values) by chromosome (negative root growth values not included)

Figure 4. Page 53
Significant markers ($p<0.01$) for NRG, RRG, and BLUP estimates (shown as –log p values) by Chromosome (negative root growth values included)

LIST OF TABLES

Chapter 1

Table 1. Page 9
Al tolerance transporter genes involved in citrate and malate efflux that have been previously identified and characterized in several species and the known regulation of expression.

Table 2. Page 13
QTL identified in maize Al tolerance studies. Locus or gene names, types of markers used in the study, and the population used are shown

Table 3. Page 14
Broad (H^2) and narrow (h^2) sense heritability estimates for Al tolerance in maize populations reported in linkage and association studies.

Chapter 2

Table 1. Page 27
Candidate Al tolerance genes evaluated by association mapping.
Thirteen genes were selected from gene expression analysis, while another eight came from comparative genomics. Genes were identified from genome survey sequence contigs created by the MAGI Project. MAGI build version and reference number are reported. ‘‘Length’’ describes the total length of sequence used for polymorphic site identification. ‘‘# Lines’’ refers to the number of entries with sufficient information to include in the association analysis. ‘‘# Sites’’ refers to polymorphisms that occurred at greater than 10% frequency. Physical-genetic map locations for each candidate gene are reported according to chromosome, genomic sequencing contig and genetic map bin. Genetic bins that appear in bold represent those under previously reported Al tolerance QTL. Gene *AL16* was evaluated by a large indel (*HAP) rather than by gene sequence. Gene *ASL* underwent two rounds of sequence analysis.

Table 2. Page 28
Evaluation of association mapping results by ANOVA.
GLM and MLM analyses were used to evaluate the 21 candidate Al tolerance genes, using the net root growth trait collected from Al treated plants. These models incorporated the population structure (Q) of the Diversity Panel, the relative kinship (K) of the Diversity Panel and net root growth of the Diversity Panel grown without Al stress as a fixed effect covariate (C). The GLM model used only factor Q, while the MLM models used factors Q+K and Q+K+C. Six candidate genes gave significant results and are shown, with the number of significant sites ($p<0.01$) identified per locus for each model. The maximum value for variance explained by a marker within a gene in any model is reported. False Discovery Rates were empirically calculated for each model based on 1,095 random SNPs throughout the genome and are expressed as

percentages. Narrow (h^2) and broad sense (H^2) heritability estimates were generated for each trait based upon variance estimates from the MLM.

Table 3. Page 30
Evaluation of linkage mapping results by ANOVA.
GLM analysis was used to evaluate whether SNP markers within candidate Al tolerance genes explained significance variance for Al tolerance observed in F2 populations. Gene action was modeled as either additive (‘‘add’’) or dominant (‘‘dom’’) based on allelic means. The variance explained by each significant SNP is reported. As both *SAHH* and *ZmASL* were significantly associated with Al tolerance for the B73xCML247 population, a summary model is reported. DF: Degrees of Freedom; SS: Sum of Squares; F: F ratio; P: P value.

Chapter 3

Table 1. Page 52
Number of markers identified in the model ($p < 0.01$) and r^2 for each phenotype, not including negative root growth

Table 2. Page 53
Number of marker effects in the model ($p < 0.01$) and r^2 for each phenotype, including negative root growth

LIST OF ABBREVIATIONS

ALMT	Aluminum Activated Malate transporter
BLUP	Best Linear Unbiased Prediction
Chr	Chromosome
CNV	Copy Number Variation
FDR	False Discovery Rate
FRG	Final Root Grown
GLM	General Linear Model
GWAS	Genome Wide Association Studies
IBM	Intermated B73 Mo17
IRG	Initial Root Growth
ISL	Isocitrate Lyase
K	Kinship matrix
LD	Linkage Disequilibrium
MATE	Multi drug And Toxic compound Extrusion
ME	NADP-Malic Enzyme
MITE	Miniature Inverted Transposable Element
MLM	Mixed Linear Model
NAM	Nested Association Mapping Population
NRG	Net Root Growth
OA	Organic Acid
PM	Plasma Membrane
PME	Pectin MethylEsterase
Q	Population Structure
QTL	Quantitative Trait Loci
RIL	Recombinant Inbred Line

ROS	Reactive Oxygen Species
RRG	Relative Root Growth
SAHH	S-adenosyl-L-Homocysteinase
SNP	Single Nucleotide Polymophism
ZmASL	Zea Mays AltSb Like

LIST OF SYMBOLS

Al	Aluminum
Ca	Calcium
Fe	Iron
H	Hydrogen
h^2	narrow sense heritability
H^2	broad sense heritability
K	Potassium
Mg	Magnesium
Mn	Manganese
Mo	Molybdenum
N	Nitrogen
P	Phosphorous

CHAPTER 1
ACID SOILS AND ALUMINUM TOXICITY

Acid Soils

Soil is a complex environment made up of organic and mineral particles, liquids, gases and a diverse range of biological organisms. Many aspects influence the soil and its pH, such as climate, parent material, topography, and biological factors [1]. Acidification of soil is due to the increased concentration of H^+ ions and is determined by mineral composition of soil, ion exchange and hydrolysis reactions [2-4]. Highly acid soils are generally considered those with pH < 5.5, cover about 30% of world's total land area and 70% of the world's potentially arable land. The American continents account for 40.9% of the world's acid soils, followed by Asia (26.4%), Africa (16.7%), Europe (9.9%), Australia and New Zealand (6.1%)[5].

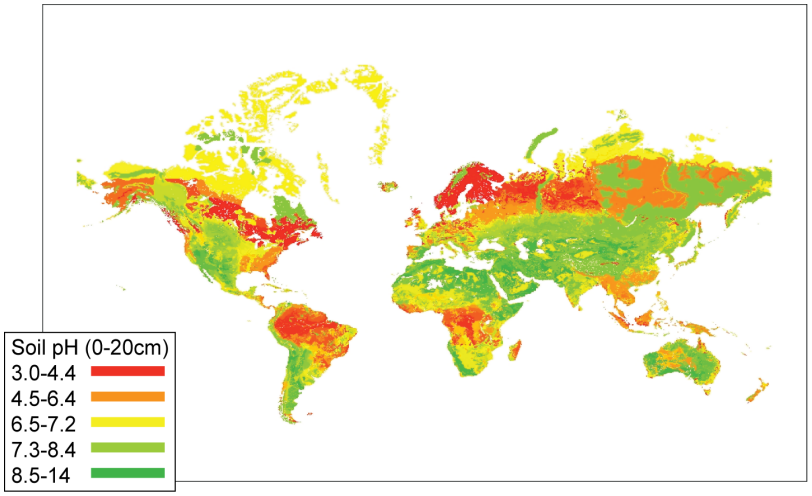

Figure 1. World Soil pH (FAO-UNESCO)

Acid soils occur in two main global belts, a northern cold temperate belt and a southern tropical belt (Figure 1). A majority of these soils are forested and provide the bulk of the world's timber. The rest is mainly savanna, prairie or steppe vegetation. The arable acid soils in the northern belt are cropped mainly to cool-season Pooid cereals (rye, oats, wheat, barley), coniferous forests cover the rest. More than half of the acidic soils found in the tropics and subtropics are classified as Oxisols or Ultisols and are mostly covered with forests or wetland. Those soils used for agriculture are cropped to tree crops, including oil palm, tea, coffee, tropical fruits, and rubber, as well as legumes, root and tuber crops and more than a third of cereals. These acidic tropical soils are considered the largest potential source for future agricultural development, but forests and wetlands of these areas are very valuable for ecosystem balance and preservation of biodiversity. Over 11 M ha of these ecosystems, mostly forests, is cleared every year. Transformation of a tropical primary forest to a forest plantation or cropland depletes much of the nutrient reserves, and a total loss of fertility occurs after only 2-3 crop rotations. Only a small portion of this land is

maintained under productive and sustainable systems, while most has become relatively unproductive [4, 5]. This creates a need to balance the conservation of these important tropical forests with increasing agriculture production by either expansion of croplands or increased production of those lands already cleared. The use of crops tolerant to these soils can contribute to that increased production.

The poor fertility of acid soils is due to a combination of mineral toxicities (Al and Mn) and deficiencies (P, Ca, Mg, Mo, and sometimes Fe) (Figure2). Acid soils usually have low water holding capacity and are susceptible to crusting, erosion and compaction. Soil microbial communities are also negatively affected [3-5]. While these factors all affect soil fertility, the most important factor restricting plant growth on acid soils is toxicity of soluble aluminum (Al). Poor crop growth can be directly correlated with Al saturation of soils [2, 5].

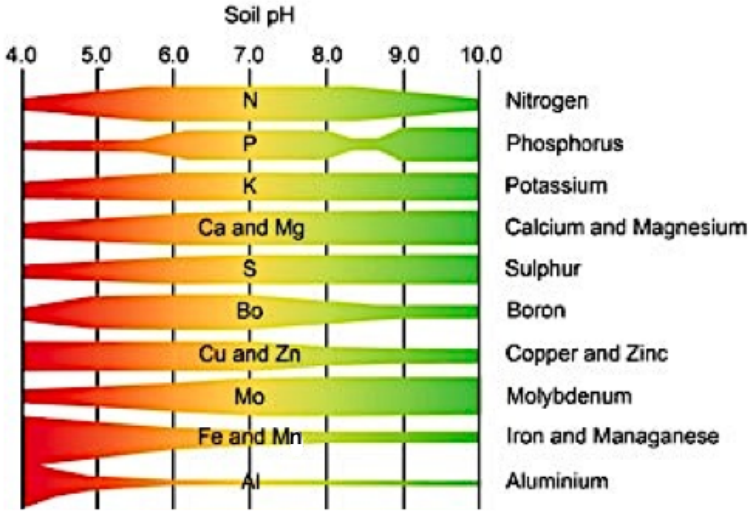

Figure 2. Effect of pH on bioavailability of minerals

Aluminum (Al) is most abundant metal and the 3rd most abundant element in the earths crust [6]. Soils contain an average of 7% Al [1]. In soils with pH >5, Al primarily occurs in insoluble minerals and complexes harmless to plants [3]. In order for Al in soil to be bioavailable it must first be in solution, which can happen through several processes [7].

There are numerous Al containing minerals and compounds in the soil that can dissolve under acidic conditions and release soluble Al into the soil solution. Soils formed from granite parent materials acidify at a faster rate than those from calcareous material. Sandy soils and highly weathered soils acidify more rapidly than clay, due to leaching of nutrient cations [1]. Al in soils forms the structure of primary and secondary minerals, such as feldspars, micas, kaolins, vermiculites, and aluminosilicates (Al_2SiO_5), a major component of clay soils. As soil weathers, silicon

and phosphates are leached away in solution, leaving Al behind in solid forms of Al hydroxides, such as gibbsite ($Al(OH)_3$), which predominates at higher or neutral pH. These Al hydroxides can dissolve under acidic conditions, leaving aqueous Al species in the soil, which also increases H^+ concentration in the soil solution. Inorganic compounds in many acid soils release Al^{3+} and Fe^{3+} ions along with other metals, enabling hydrolysis (splitting of H-O bonds in water). These reactions produces several different species of soluble Al (Figure 3) but Al^{3+} ($Al(H_20)_6^{3+}$) has been established as the most phytotoxic form of Al below pH 5. Al^{3+} is a reactive metal that can form compounds with organic and inorganic compounds, such as carboxylate, phosphates, sulfates, and organic acids forming less toxic complexes [1, 3-5, 7]. The measurement of Al^{3+} activity is more closely related to Al^{3+} toxicity than soluble Al concentrations. Al^{3+} interacts with other soil components that can alter ionic strength, availability of binding sites, surface potential of charges surfaces and form harmless complexes, which may change or reduce the potential of Al^{3+} to injure plants [3, 8] No evidence suggests Al is essential for plants or animals [1].

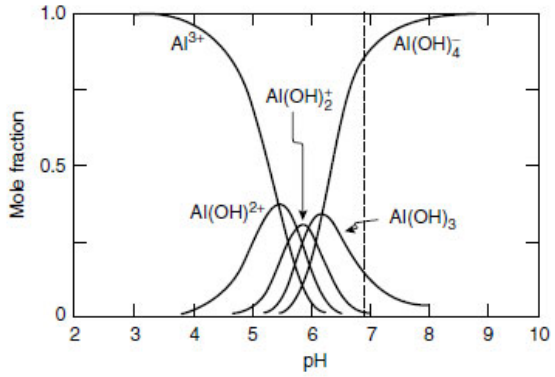

Figure 3. Effect of pH on aluminum species

Soil acidification can be intensified through and natural processes, such as acid rain, nitrification, and microorganism metabolism, as well as management practices. Intensively cropped soils, excessive application of ammonium-based fertilizers and long-term cultivation of legume-based pasture systems can all increase soil acidity, because all these processes lead to excess nitrate in the soil. Excess N, either from nitrification (conversion of ammonium to nitrate, which releases H^+ ions), or from N fixing plants, leads to increased risk of nitrate leaching taking basic cations (K^+, Ca^{2+}, Mg^{2+} and Na^+) with it [2, 3, 5]. The excess H^+ ions of acid soils replace these cations, which are normally attached to particles in the soil, making them available to be leached away. When negatively charged nitrate (NO_3^-) is leached from the soil it takes these positively charged cations with it to maintain electrical balance. Continued removal of vegetation or harvesting of crops, also removes important nutrient cations due to increased leaching and run-off during inadequate vegetation cover. Management practices contributing to rapid increase in soil acidification have become a problem in areas of Australia, where 1/3 of the soils are acidic, and China, where population pressure has increased demands for food.

A combination of Al tolerant varieties, sustainable management practices, proper use of fertilizers and amending of acid soils would be an efficient way to improve crop yields and fertility of these acid soils. Amending or liming with substances such as calcium carbonate ($CaCO_3$) raises pH, returns basic cations to the soil, and can precipitate out soluble Al^{3+} species, but does not improve subsoil acidity. This strategy is often not an option for resource poor farmers who instead apply locally available compost and rely on crops adapted to acid soils, even though there is high demand for less tolerant crops, such as maize, in these areas [2, 3]. Solid organic matter can retain Al by adsorption or complexation. A 40% reduction in soluble Al was reported after addition of 2% decomposed leafy material [1, 9].

Effects and mechanisms of Al toxicity

The earliest response to Al toxicity is the rapid inhibition of root growth. This is followed by other root stress responses, such as disruption of the root cap, callose formation, lignin deposition, reduced cell division, increased root diameter, and induction of lesions in epidermal and cortical tissues near the elongation zone [10]. Although Al can affect both cell elongation and cell division, the rapidity of the response suggests that cell elongation is primarily affected, because the timing of the root cell cycle is such that it takes approximately 24 hrs for a cell to divide. The root apex is a major site of Al-perception and induction of response mechanisms. The distal transition zone, localized to first 2-3 mm of root between meristem and elongation zones, has been identified as the most Al-sensitive part of the root in several species. This is the area where cells switch from division to elongation, and essentially where growth is occurring. The root cap was previously thought to play a role in Al-perception in the root, but studies found the removal of the root cap produced no change in root growth under Al toxic conditions [1, 3, 7, 10, 11].

Prolonged exposure to toxic Al causes the primary roots to exhibit an abnormal morphology, becoming short, stubby, thickened and brittle, with little or no fine branching, along with inhibited lateral root formation. This ultimately affects whole plant biomass and crop yields by restricting the absorption of water and nutrients (Ca, Mg and K). The plants become short and stunted, with reduced the number of ears per plant and delayed flowing. The leaves become small and dark often with chlorotic patches and marginal necrosis, indicating suppressed photosynthesis. Inhibition of symbiosis with rhizobia has also been documented [1-3, 6].

The physiological mechanisms for Al tolerance have been more extensively studied than the molecular bases, but neither is well understood. Several physiological mechanisms have been proposed but only a few have been experimentally verified. Many physiological and molecular mechanisms for Al tolerance may exist among and even within species. Al toxicity, like other abiotic stresses, creates a network of interconnected primary and secondary stress responses; therefore identifying a single mechanism may not be possible. Since inhibition of root growth at the root apex has

been well documented as the most immediate response of Al toxicity, much research has been focused on this response [3, 7, 11].

It is still unclear if Al initiates stress in the apoplast or symplast. The apoplast (outside the plasma membrane) is a logical primary site for Al stress because Al ions from soil solution can reach this area without crossing the membrane barrier. Highly charged cations like Al^{3+} bind strongly to fixed negative charges within the cell wall and on the plasma membrane (PM) surface. The root cell wall accumulates the majority of the root-associated Al (approximately 90%) and other cations in major crop species. Higher accumulation in the root cell wall has been correlated with increased Al sensitivity. In a study of several plant species differing in tolerance, the Al resistant genotypes had less negative membrane surface potential than sensitive genotypes [1, 7].

In cell walls, Al can bind to pectin and modify synthesis or deposition of polysaccharides. A major function of pectin is to provide charged structures for ion exchange in cell walls. Binding of Al to pectin can stiffen cell walls, inhibit cell elongation, and restrict flow of solutes and macromolecules through the apoplast. Al could also disrupt normal cell wall growth by reducing Ca concentration below that required for cross-linking of pectin residues or displacing Ca from root cortical cell walls [1, 3].

In the PM Al can bind to negatively charged phospholipids, displacing other cations in the membrane and interfere with protons involved in transport and signal transduction processes. Al is reported to interfere with the uptake of many nutrients and water by blocking membrane-bound transport proteins or ion channels, especially those involved in Ca and K influx [1, 3, 4, 12].

Only a fraction of Al enters the symplasm (inside the plasma membrane), as it is a trivalent cation and cannot easily cross the PM. It enters the cytoplasm through endocytosis or nonspecific transporters. Once in the symplasm, Al can interfere with many metabolic processes and cellular functions, including the cytoskeletal function, calcium homeostasis, phosphorus metabolism and phytohormone function. Due to its higher affinity, Al outcompetes other cations such as Mg and Ca for important binding sites and can bind with internal membranes in chloroplasts or nuclei, reducing photosynthetic activity and basic cellular functions. Al causes oxidative stress, generating reactive oxygen species (ROS), which trigger induction of several proteins to combat oxidative stress such as glutathione S-transferase. Inside the cell, Al can also form complexes with organic ligands, reducing its toxicity. This is one internal mechanism plants use to tolerate Al stress [1, 3].

Al tolerance/resistance mechanisms in plants

Plants vary widely in their ability to tolerate Al toxicity primarily through two main mechanisms, Al exclusion and internal tolerance (Figure 5). Exclusion

mechanisms prevent Al from entering the plant, while tolerance mechanisms allow the plant to take up Al and store it safely within the plant [4, 10]. Internal tolerance mechanisms are often seen in plant species endemic to regions with acidic soils. Some species even thrive under these conditions [4]. In this type of mechanism, Al entering the symplasm is bound by internal ligands, such as organic acids or flavanoids, into harmless complexes (Figure 5A). These complexes are then stored in vacuoles or other organelles and can be transported to other regions of the plant such as the leaves. Plants such as tea (*Camelia sinensis*) and buckwheat (*Fagopyrum esculentum*), which are highly tolerant to Al, can store high amounts of these types of Al compounds in their leaves. This type of internal tolerance mechanism is usually accompanied by rapidly induced of genes involved in cell wall structure and ROS detoxification to quickly repair cell wall damage and minimize oxidative stress [1, 3].

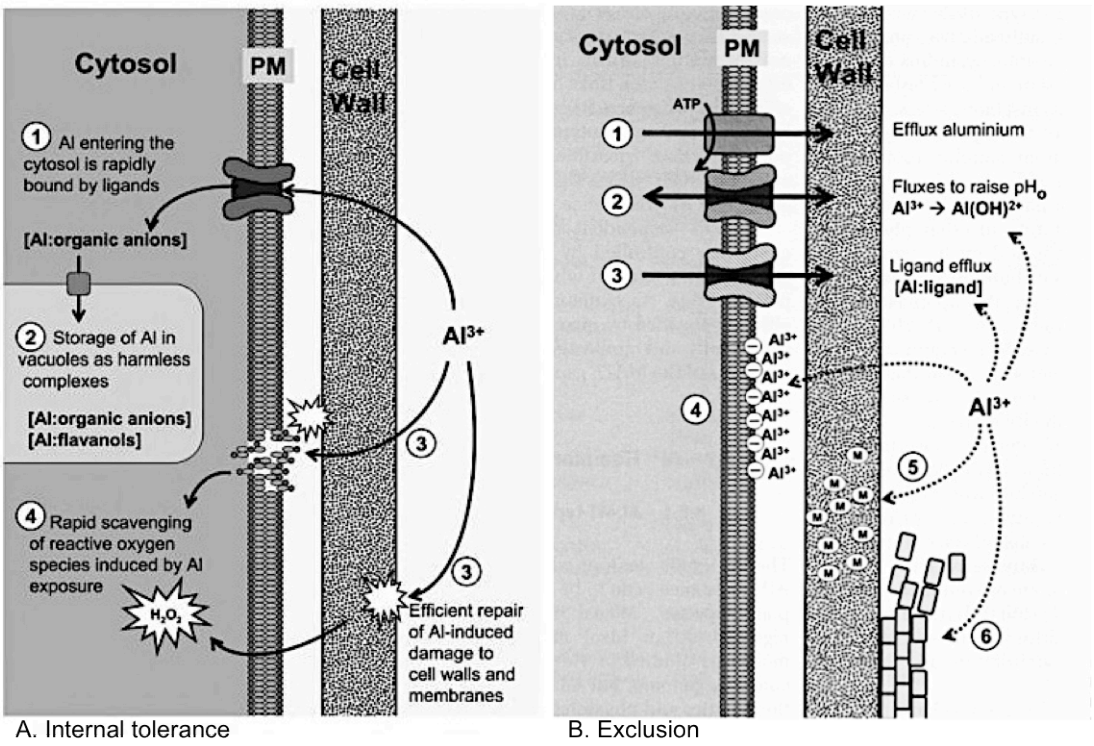

Figure 4. (A) Internal and (B) Exclusion mechanisms for Al^{3+} tolerance in plants. Plasma membrane (PM). **A. Internal tolerance.** (1) Binding of Al by ligands to form harmless complexes. (2) Storage of these complexes in vacuoles or other organelles. (3) Rapid repair to the cell wall and membranes. (4) Scavenging of reactive oxygen species (ROS). **B. Exclusion mechanisms.** (1) Lowering Al concentration in cytosol, by active export or vesicular transport. (2) Transport processes across PM that raise apoplastic pH and reduce molar fraction of Al (3) Efflux of ligands that bind with Al to form harmless complexes (4) Lowered charge density on membrane surface reducing Al binding sites (5) Increased pectin methylation reduces binding of Al to cell wall (6) Growth and detachment of border cells [3].

Exclusion mechanisms function to minimize the interaction of Al with the root apoplast and prevent Al from entering the symplast (Figure 5B). Several of these mechanisms use membrane transporters, which either decrease Al concentrations in cytosol or transport compounds, such as organic acids (OA), phosphate, polypeptides, or phenolics, to alter the pH or bind with Al in the apoplast or rhizosphere. Before Al can reach the root apical meristem it must cross the mucilage, which is a complex mixture of polysaccharides, border cells and cell wall fragments produced by root cap that may contribute to Al tolerance. Differential binding of Al to mucilage, the cell wall, and the outer surface of the plasma membrane can decrease the accumulation of Al in the root symplasm. Changes in charge density of the PM or pectin methylation of the cell wall can reduce the number of negatively charged binding sites for Al. Al accumulation in sensitive varieties of maize, rice and common bean has been correlated with lower cell wall pectin methylation. Some species, such as buckwheat, have border cells which grow and detach, taking Al bound in those cells with it. Interaction with mycorrhizal fungi has also been proposed to play a role in conferring tolerance [1-3, 7, 11, 13].

Root tip organic acid (OA) exudation is the most extensively studied mechanism of Al exclusion in plants. There is extensive physiological and genetic evidence linking OA exudation (citrate, malate and oxalate) to Al tolerance in several plant families and across several species [10]. Organic anions bind Al in the rhizosphere to form nontoxic complexes, which help reduce Al accumulation in root cell wall. Most species exude either malate or citrate, while some excrete both, and a few species exude oxalate. Citrate (a tricarboxylic acid) is more effective at chelating Al than malate or oxalate (dicarboxylic acids) due to its stronger affinity for the Al^{3+} ion. Many Al tolerant genotypes release large amounts of OA from roots and also have less Al accumulation in root tip (apoplasm and symplasm). However, some sensitive varieties have high OA efflux without the benefit of increased tolerance [14]. Rice, one of the most Al tolerant cereals, shows no apparent correlation between root OA release and Al tolerance suggesting that mechanisms other than root OA exudation are used to achieve its high level of tolerance [3, 10, 15, 16].

In species that use OA exudation as an Al tolerance mechanism, the location, type, timing and amount of OA release may differ. The first and most detailed study on OA efflux from roots as an Al exclusion mechanism was in wheat. Wheat OA efflux occurs in a highly localized portion of root apex, but in maize it is released from much broader region encompassing the root tip. Barley exhibits differences in the root location of citrate release that is dependent on expression differences of an Al tolerance gene [10, 17]. Some species show a rapid OA efflux response (wheat/malate), while others (sorghum/citrate) exhibit a lag. In sorghum, a 6 day exposure of roots of tolerant genotypes to Al is required to fully induce Al tolerance, citrate exudation and the expression of the major Al tolerance gene, *SbMATE* [18]. Citrate is released in rye is increased at a steady rate over time. These differences

suggest various mechanisms are involved in Al tolerance between and within species [3, 10].

A rapid response of OA release suggests the mechanism is occurring at the protein level, while a lag in response time is more indicative of a transcriptional or translational response. Membrane transporters in several species have been implicated in the OA efflux response, but so far there has been no direct evidence of changes in OA biosynthesis, although overexpression of OA metabolic enzymes in some species has enhanced tolerance [10, 19] .

Genetics of Al tolerance in plants

Plants vary widely in their ability to tolerate acid soils. Effects of Al toxicity in 22 species over 7 plant families, including Poaceae, show several species or genotypes within species are resistant or tolerant to Al [1, 6]. Rice is 6-10 times more Al tolerant than other cereals, followed by rye, triticale, maize, sorghum, wheat, and barley. Species endemic to regions with acidic soils usually have some level of Al resistance [2, 6, 16, 20]. This genetic diversity provides a valuable resource that breeders can use to develop more resistant varieties.

A large portion of food for humans and animals comes from cereal crops (Poaceae), so this family has been a primary focus of Al tolerance research. The abundant genetic resources and large research communities for these crops make them ideal model species for genetic and physiological studies. Although considerable diversity exists within and between these species there is conservation in gene content, making comparative genomics studies a useful tool [21].

Genes controlling Al resistance have been identified by use of segregating populations and mutants, including wheat, rye, rice, *Arabidopsis*, barley, sorghum and maize. In wheat, sorghum and barley, Al tolerance is inherited as a simple trait controlled by one or a few major genes. In maize, rice, rye and *Arabidopsis* tolerance is quantitatively inherited [1].

A majority of Al tolerance genes identified in several species come from two distinct but unrelated gene families that encode for transport proteins involved in Al activated organic acid release, specifically citrate and malate. Variation for Al tolerance is correlated with expression of these genes in root apices. Al tolerance genes from the aluminum-activated malate transporter (ALMT) and multi-drug and toxic compound extrusion (MATE) families of transporter proteins have been identified in several species (Table 1).

Species	Gene	transport	Reference	regulation of expression
Wheat	*TaALMT1*	Malate	(Sasaki 2004)	tandem repeats in promotor
Wheat	*TaMATE1*	Citrate	(Ryan 2009)	Transposon upstream ORF
Sorghum	*SbMATE*	Citrate	(Magalhaes 2007)	MITE transposon
Rye	*ScALMT1 cluster*	Malate	(Collins 2008)	tandem repeats of genes
Rye	*ScMATE2*	Citrate	(Yokosho 2010)	
Maize	*ZmMATE1*	Citrate	(Maron 2010)	copy number variation
Brassica napus	*BnALMT1/BnALMT2*	Malate	(Ligaba 2006)	
Barley	*HvAACT1*	Citrate	(Furukawa 2007)	transposon upstream of ORF
Arabidopsis	*AtALMT1*	Malate	(Hoekenga 2006)	*STOP1*
Arabidopsis	*AtMATE1*	Citrate	(Lui 2009)	*STOP1*

Table 1. Al tolerance transporter genes involved in citrate and malate efflux that have been previously identified and characterized in several species and the known regulation of expression.

ALMT genes

ALMT-type proteins are a small family unique to plants, which mediate passive transport of both organic and inorganic anions. They encode anion channels but differ in substrate specificity, membrane location, and function. ALMT genes have been characterized for several species but only a few are known to confer Al tolerance. Members of the ALMT family appear to be involved in several physiological processes including Al tolerance, mineral nutrition and guard cell regulation [22, 23].

The first Al tolerance gene isolated, cloned, and characterized from any plant species was an ALMT gene, *TaALMT1,* in wheat (*Triticum aestivum*). *TaALMT1* encodes a plasma membrane transporter protein constitutively expressed in the root tips of Al tolerant genotypes and facilitates malate efflux when exposed to Al. Tolerant genotypes accumulate 3-8 fold less Al in the root apex than sensitive ones [24]. This gene co-localizes with a previously identified locus on chromosome 4DL, *Alt1*, which co-segregates with Al activated malate efflux from roots. Although 3 major QTLs for resistance in wheat have been identified, this one locus explains ~70% of genetic variation and is considered the major Al tolerance locus in wheat [17, 24]. *TaALMT1* was initially isolated using near-isogenic lines of wheat that vary in Al resistant at a single locus, but was later linked to variation of Al tolerance in several wheat mapping populations and across 13 wheat cultivars. Heterologous expression of *TaALMT1* in transgenic rice, tobacco, barley, wheat, and Arabidopsis has also been shown to confer Al resistance [25, 26].

At least two *TaALMT1* alleles are present in wheat; however, Al tolerance is correlated with level of expression and not through different alleles. Expression correlates with the number of tandem repeats in a region upstream promoter region of *TaALMT1* (triplications enhance expression). Although both gene expression of *TaALMT1* and malate efflux are strongly correlated with Al tolerance, there is a poor

correlation between *TaALMT1* expression and malate efflux. This suggests other genes contributing to malate efflux or tolerance are involved [17, 24]. In wheat both sensitive and tolerant lines have the *TaALMT1* transporter, but there is much higher constitutive expression of *TaALMT1* in the tolerant genotypes [10]. *TaALMT1* expression is not upregulated by Al, but the OA transporting activity of the protein appears to be activated by the presence of Al [24].

A *TaALMT1* homolog in *Arabidopsis thaliana, AtALMT1,* encodes an Al activated root malate transporter that mediates malate exudation in roots and controls a majority of Al tolerance. Al tolerance in *Arabidopsis* is a complex trait that uses malate efflux as a primary mechanism, but only 70% of the variation is associated with malate release, suggesting one or more additional mechanisms are also involved in *Arabidopsis* tolerance [27].

Rye *(Secale cereal*) is one of the most Al tolerant species in the Triticeae. Al tolerant genotypes release more malate and citrate from roots than sensitive genotypes. Genes in rye can be a potential source of Al tolerance for wheat, as rye genes in a wheat background have been shown to confer Al tolerance. The *Alt4* locus in rye contains a cluster of genes (*ScALMT1* cluster) homologous to the *TaALMT1* gene in wheat. These are clusters of tandemly repeated genes, which differ in copy number, expression level and coding sequence. Several of these *ScALMT* genes are expressed in root tips, upregulated by Al and expression levels are correlated with Al tolerance [28]. However, it is likely that some ALMT genes in the cluster could be involved in other transport functions unrelated to Al tolerance, such as those found for ALMT homologs in other species (*HvALMT* in barley*, AtALMT9* and *AtALMT12* in *Arabidopsis*, *ZmALMT1* and *ZmALMT9* maize) [22, 23, 28-30].

Oilseed rape *(Brassica napus)* was found to have Al activated release of malate and citrate from roots. Two *TaALMT1* homologs *BnALMT1* and *BnALMT2* encode Al induced malate transporters and are expressed in roots. Expression in transgenic tobacco increased Al resistance, but they have not yet been confirmed as contributing to Al resistance [4, 31]

MATE genes

The MATE (multi-drug and toxic compound extrusion) family of proteins is a large and diverse group found in all living organisms. In plants there are a large number of genes encoding MATE proteins suggesting an array of biological functions. Only a few have been functionally characterized as secondary transporters, exporting a wide variety of substrates and organic compounds. The first MATE genes functioned to export cationic drugs and toxic compounds, hence their name. Several plant MATE proteins have recently been found to mediate citrate transport, iron transport, and other secondary metabolites. Others have specificity for substrates such as flavonoids, xenobiotics, and alkaloids.[32]

The subset of MATEs that transport citrate are involved with Al tolerance, Fe nutrition or P efficiency. MATE genes have been identified as major Al tolerance genes in sorghum (*SbMATE*) barley (*HvAACT1*) and maize (*ZmMATE1*) [18, 33, 34]. MATEs have also been identified or implicated in minor QTL or secondary tolerance in wheat (*TaMATE1)*, rye (*ScMATE2)*, maize (*ZmASL*) and Arabidopsis (*AtMATE1)* [19, 25, 35, 36].

The first MATE gene involved in Al resistance, *SbMATE*, was positionally cloned in sorghum *(Sorghum bicolor)*. It underlies a major Al tolerance locus, Alt_{SB}, which explains 80% of the variation in the mapping population and can confer up to a ten-fold increase in tolerance. *SbMATE* encodes a transport protein located on the root tip cell plasma membrane that facilitates citrate release from root cells. It is constitutively expressed in roots of tolerant lines and barely detectable in sensitive ones, but Al treatment significantly increases its expression, which correlates with Al induction of citrate release and Al tolerance [4, 18].

Coding regions of *SbMATE* are identical in sensitive and tolerance genotypes. Data from genetically diverse sorghum lines indicate mutations in regulatory regions enhance gene expression in root apex. A miniature inverted transposable element (MITE) was found in a repeated region of the promoter. MITEs identified in noncoding regions of genes can alter gene expression[25]. Variation in size of this region, due to copy number variation (CNV) of the MITE, significantly correlates with expression of SbMATE and Al tolerance, however it does not explain all of the Al tolerance variation of this gene. Although a minimum MITE insertion size in the promoter region is needed for Al tolerance, it may be in LD with or interact with other polymorphisms. Within the significant genetic diversity seen in sorghum, there may be other Al tolerance genes or mechanisms contributing to the wide phenotypic variation for Al tolerance [18, 25].

SbMATE homologs controlling citrate efflux have also been found in barley, *Arabidopsis*, maize, wheat, and rye. Maize genes will be discussed separately in the following section. The *HvAACT1* gene (*HvMATE*) in barley (*Hordeum vulgare L*) is a homolog of *SbMATE* that encodes a citrate transporter and enables its adaptation to acid soils. Similar to *TaALMT1*, *HvAACT1* is constitutively expressed in root apices but requires Al to active citrate release, with greater expression seen in tolerant varieties. A transposon upstream of coding region enhances and alters the location of expression in roots from the root xylem parenchyma where it is presumed to play a role in xylem Fe translocation, to the root tip epidermis [4, 25].

A MATE transporter found in wheat, *TaMATE1*, thought to mediate root citrate release is a possible secondary mechanism to confer Al tolerance. The *TaMATE1* protein in wheat is not activated by Al, but confers constitutive citrate efflux from root apices. So far this trait has only been found in a few very tolerant cultivars from Brazil. Expression of *TaMATE1* is associated with a transposable element upstream of the coding region [4, 25, 36].

AtMATE in *Arabidobsis*, a close homolog of *SbMATE*, was found through mutant analysis and confers citrate release. It is a minor Al tolerance mechanism that functions with *AtALMT1* malate release to confer the full range of *Arabidopsis* Al tolerance. Al tolerance in *Arabidopsis* was not associated with increased expression of genes encoding OA synthesis but with differential expression of transporters [37]. The transcription factor, *STOP1*, regulates expression of both *AtALMT1* and *AtMATE1* [35]. Mutational analysis also identified *AtALS1* and *AtALS3* genes which encode 'half' ABC transporters required for Al tolerance by redistributing Al away from sensitive tissues [25, 38].

The underlying mechanisms by which these transport genes are activated by Al is still not understood, although there is evidence for the implication of basipetal auxin transport in signaling [8]. For some genes Al activates the transport function of a preexisting protein (*TaALMT1* and *HvAACT1*), while for other genes Al induces expression and then also activates transport activity (*SbMATE*) [3, 10, 25].

QTL studies of Al tolerance in maize

Maize is the most widely grown crop in the world. It is one of the most important food sources for both humans and animals and is becoming increasingly important for fuel and ethanol production. Maize is widely adaptable and highly productive in a variety of climates and environmental conditions, due to its genetic diversity, used of hybrids and improved crop management practices. In 1995 Approximately 20% of all maize was grown on acid soils, with up to a 70% reduction in yields due to Al toxicity [5]. There is considerable variability for Al tolerance in maize, which has been exploited by breeders. High yielding elite germplasm is generally sensitive to Al and cultivars from the South American regions with high Al tolerance generally have lower yields than elite germplasm [39]. Identification of tolerant alleles may be used in marker assisted breeding to incorporate them into this elite germplasm, creating higher yielding Al tolerant varieties.

While many studies have contributed to the physiological and genetic knowledge of Al tolerance in maize, the molecular basis for tolerance is still poorly understood. Al tolerance is correlated with reduced Al accumulation in the root tip so the focus has been on exclusion mechanisms[15]. Organic acid (OA) release is believed to be the main mechanism of Al tolerance in maize because Al tolerant genotypes exhibit an Al activated citrate release and in some cases smaller amounts of malate. However, the correlation between Al resistance and Al activated citrate release does not hold up when looking at a more diverse selection of germplasm [14]. There was a high correlation between Al tolerance and Al root tip exclusion, but not between Al tolerance and citrate release. Some Al sensitive lines also exhibited large amounts of citrate release [14, 15, 37, 40].

QTL studies have identified major and minor loci contributing to Al tolerance in a variety of populations (Table 2). All of these QTL studies measure Al tolerance

as root growth in an Al^{3+} toxic nutrient solution [41]. Different crosses have produced differing results in the number and location of major loci, which is expected when using in different linkage populations with different alleles or coverage of markers [42-44].

BIN	locus/gene	Marker	lines	References
2		RFLP	S. American lines	Brondani and Paiva 1996
2.06	QTL1	RFLP/SSR	L53 x Al327	Ninamango-Cardenas 2003, Mattiello 2012
3.04		SSR	IBM (B73 x Mo17)	Hoekenga/Mason 2005*
3.07		SSR	IBM (B73 x Mo17)	Hoekenga/Mason 2005*
4.01		SSR	IBM (B73 x Mo17)	Hoekenga/Mason 2005
4.03		SSR	diallele (Tuxpeño/Cateto origin)	Conceição 2009
5.01		SSR	diallele (Tuxpeño/Cateto origin)	Conceição 2009
5.02	*ZmMATE2*	RFLP/SSR	L32 x Al327	Guimaraes 2009*, Maron 2010
6.00	QTL2	RFLP/SSR	L53 x Al327	Ninamango-Cardenas 2003, Mattiello 2012
6.00	*ZmMATE1*	RFLP/ SSR	L32 x Al327	Guimaraes 2009*, Maron 2010
6.01	*Alm2*	RFLP	Cat-100-6 x S1587-17*	Sibov 1999, Mattiello 2012
6.05	QTL3	RFLP/SSR	L53 x Al327	Ninamango-Cardenas 2003, Mattiello 2012
6.05		SSR	diallele (Tuxpeño/ Cateto origin)	Conceição 2009
8		RFLP	L53 x Al327	Torres 1997
8.03		SSR	IBM (B73 x Mo17)	Hoekenga/Mason 2005*
8.04	QTL4	RFLP/SSR	L53 x Al327	Ninamango-Cardenas 2003, Mattiello 2012
8.05		SSR	diallele (Tuxpeño/ Cateto origin)	Conceição 2009
8.07	QTL5	RFLP/SSR	L53 x Al327	Ninamango-Cardenas 2003, Mattiello 2012
10.00		SSR	IBM (B73 x Mo17)	Hoekenga/Mason 2005*
10		RFLP	Cat-100-6 x S1587-17*	Moon 1997
10.01		SSR	diallele (Tuxpeño/Cateto origin)	Conceição 2009
10.03	*Alm1*	RFLP	Cat-100-6 x S1587-17*	Sibov 1999, Mattiello 2012
10.03		SSR	IBM (B73 x Mo17)	Hoekenga/Mason 2005*
			*somaclone	*unpublished
			Al237 / L1327 / Cateto Al 237/67	

Table 2. QTL identified in maize Al tolerance studies. Locus or gene names, types of markers used in the study, and the population used are shown.

The earliest studies used restriction fragment length polymorphisms (RFLP) in segregating F_2 populations. These studies report Al tolerance as a monogenic dominant trait controlled at a single locus by a multiple allelic series. Major QTL were found on regions of chromosome 2, 6, 8 or 10, and are rather broad, because they are based on recombination maps and populations with little recombination [39, 43, 45-47].

With advances in genetic resources, QTL studies now report maize Al tolerance as a complex quantitatively inherited trait with additive gene effects from multiple loci or tolerance mechanisms. Most QTLs differ but have overlapping regions. Two regions identified in multiple studies are located on the short arms of chromosomes (Chr) 10 and 6, referred to as *Alm1* and *Alm2*. A major Al tolerance gene *ZmMATE1*, was found underlying the *Alm2* locus [33]. Sibov et al (1999) concluded that Al tolerance is controlled by the recessive epistastic interaction of these two loci, showing a 9:3:4 ration in a F_2 population. Chr 8 loci are also reported in several studies [10, 39, 42-44, 48, 49].

Most of these studies use parents that originate or are progeny of lines from South American areas with highly acidic soils, with the exception on one study using North American lines of an intermated recombinant inbred population (IBM), whose parents have a similar level of tolerance. South American lines with originating from Cateto or Tuxpeno inbred lines (Cat100-6, L1327/Al237). are some of the most tolerant varieties and have been used by breeders and geneticists for many years. When these lines are paired with a sensitive line the wide range of tolerance seen in maize is represented in the resulting population. The parents of the IBM population (B73 x Mo17) have a similar level of tolerance (moderately sensitive), but show a considerable amount of transgressive segregation, suggesting additive gene effects contributing to Al tolerance from these lines. The more sensitive Mo17 line has higher Al activate citrate release rates than the most tolerant South American lines of Cateto origin. QTL mapping for Al tolerance using the IBM population identified 6 QTL accounting for about 60% of the variation in the population. Three of these QTL were donated from Mo17 and three from B73. Three QTL donated by Mo17 are epistatic and all are required for increased tolerance, but act additively to exclude Al from the root tips. Three B73 QTL are independent and unrelated to Al exclusion and act additively to increase tolerance [15, 44].

Table 3. Broad (H^2) and narrow (h^2) sense heritability estimates for Al tolerance in maize populations reported in linkage and association studies.

Reference	H^2(family means)	h^2(within family)
Ninamango-Cardenas (2003) [42]	97%	30%
Boni (2009)[50]	86%	49%
Conceição -2009 [48]	90%	
Krill 2010 [19]	41%	30%

Expression/physiology studies of Al^{3+} tolerance in maize

The advent of cDNA arrays and transcript profiling helped identify a greater number of genes whose expression is affected by Al toxicity [37]. Most of the Al tolerance genes to date (Table 1) are all Al-activated plasma membrane transporters of organic acids. The expression levels of these transport genes are highly correlated with overall Al tolerance. There is little change in the expression levels of enzymes involved in organic acid synthesis or metabolism [51]. Therefore, looking at expression levels between Al tolerant and sensitive varieties has helped narrow down prospective genes for further study.

Maron et al identified several genes of interest differentially regulated between an Al tolerant and sensitive maize lines [51]. The Al sensitive genotype accumulated significantly more Al in the root tip, indicating exclusion as consistent with a major Al tolerance mechanism. The tolerant variety showed a greater increase in the early up regulation of genes compared to the sensitive. However, over time, the number of genes up and down regulated by Al in the sensitive line increased and was ultimately greater than the tolerant line. A number of stress response genes were up regulated,

while a large number of involved in transcription, translation and ribosomal proteins were down regulated, suggesting the contribution of inhibition of cell growth and division. Several of these genes were chosen to be used in an association study (Chapter 2) [51].

Almost all studies investigating QTL or gene responses in maize measure Al tolerance as root growth in an Al toxic hydroponic solution. Mattiello et al looked at the transcript profile of maize plants grown on acid soils to more adequately mimic the soil environment of the fields on which the plants are grown [37]. This experiment can examine the role that complex soil interactions have on tolerance, such as microorganisms, border cells, and mucilage, all of which have been proposed as potential mechanisms to help plants avoid Al toxicity A hydroponic evaluation was used to determine that Al and not pH was the main expression induction factor in the soil treatment, where they are inseparable.

Differential expression patterns between the tolerant and sensitive lines clearly reflect both direct and indirect effects of stress cause by acid soil. This study was compared with that done by Maron et al [51] in hydroponics using similar genotypes. There was a minor overlap of expression results between the two growth systems, but many new physiological and transcript responses were identified. However, a variety of other variables (chip platform, lab conditions, etc) could also contribute to those differences. There was a smaller fraction of genes expressed in the sensitive line vs tolerant line, similar to Maron, indicating quick response to Al toxicity in tolerant lines prevents detrimental stress effects over time. [37]

Differential expression in several pathways related to tolerance mechanisms was observed in both expression studies. There were little changes seen in the expression of genes involved in OA synthesis, with the exception of citrate synthase in the soil treatment. However, there was an increase expression in some membrane transporters, similar to those responsible for Al tolerance in other species. 14 putative ALMT1-like sequences were identified, but determined as unlikely to play a role in Al tolerance based on expression patterns. 45 MATE-like sequences identified and some displayed patterns of Al responsive expression [37, 51].

Cell wall genes have been hypothesized as having a role in Al tolerance, since a majority of Al absorbed by the root can be localized to the apoplast. Several cell wall genes were identified in both treatments, including pectin methyesterase (PME), which was included in an association study (Chapter 2). Higher levels of PME, such as those seen in the sensitive line, might result in a higher capacity of Al to accumulate in cell walls as pectin contributes to increased cell wall negativity due to its carboxylic acid residues. Methylation of these carboxylic acid groups via PMEs can reduce the negative charges in the cell wall. Differences in pectin content and methylation of cell walls has been linked to Al tolerance and exclusion differences in maize, rice and buckwheat [13, 52, 53]. Lignin biosynthesis enzymes, also involved in the cell wall, were differentially expressed between the two lines, in both treatments [37, 51].

Al toxicity can cause oxidative stress inside the cell, producing reactive oxygen species (ROS). Several oxidative response genes, such as *glutathione S transferase* (*GST*) and *SAH hydrolase* (*SAHH*), as well as ROS genes such as *oxidase* were upregulated in the sensitive lines but not the tolerant lines, in both studies. SAHH and oxidase were also used in an association study (Chapter 2). This suggests response mechanisms in tolerant lines may act before oxidative stress occurs. However, a number of these oxidative stress genes may be a general stress response and not specific to Al toxicity. [37, 51].

Identifying genes located under previously reported QTL narrowed down results from the massive amount of microarray data. Mattiello mapped 44 genes to regions previously identified, (Table 1) on Chr 2(QTL1), Chr 6 (Alm2, QTL2, QTL3), Chr 8 (QTL4, QTL5) and Chr 10 (Alm1). Maron et al, focused on two MATE genes, *ZmMATE1* and *ZmMATE2*, located under major QTL from a population that used the same lines as the expression study [37, 51].

Al^{3+} tolerance genes in maize

Two ALMT1 homologs have been extensively studied in maize, *ZmALMT1* and *ZmALMT2*. Both are homologous to wheat *TaALMT1* and Arabidopsis AtALMT1, which encode functional Al activated transporters that confer Al tolerance. Both genes are located at bin 10.04, which does not fall under any reported major QTL (Table 2). *ZmALMT1* was closest in sequence similarity to these. *ZmALMT1* was cloned using two cultivars that differed in rates of Al activated organic acid release and tolerance. This gene has been identified as an anion transporter, but did not appear to be involved in maize Al tolerance. It most likely involved in mineral anion nutrition. [22]

ZmALMT2 was identified as a candidate for potential role in Al tolerance based on joint linkage and association analysis (Chapter 2). Out of the 7 ALMT members evaluated in this study only *ZmALMT2* was significantly associated with Al tolerance in both association and linkage analysis [19]. *ZmALMT2* shares significant homology with *ZmALMT1*, *HvALMT*, *TaALMT1*. *ZmALMT2*s potential role in Al tolerance was further examined through physiological studies. This gene was detected in both roots and shoots, with higher expression in mature regions of roots. It is localized to the plasma membrane and is constitutively expressed primarily as root malate efflux transporter. In heterologous systems (*Xenopus* oocytes and *Arabidopsis*) it was found to mediate malate and citrate transport as well as other relevant anions (Cl^- and NO^{3-}). Expression levels were higher in sensitive line (B73), which is consistent with linkage data, where the B73 allele is dominant and confers a slight increase root growth [19]. However, *ZmALMT2* expression was not activated or enhanced by the presence of Al^{3+}, but was suppressed, suggesting a limited role in Al tolerance. Like *ZmALMT1*, *ZmALMT2* is also believed to play a role in mineral nutrition acquisition and transport,

but further studies are needed to rule out its role in Al tolerance. Further discussion on *ZmALMT2* can be found in Chapter 3, conclusions [23].

Maron et al [32] used a combination of genetic and functional approaches to identify and characterize two genes of MATE family *ZmMATE1*and *ZmMATE2*, underlying two major QTL on Chr 5 and 6 found in several mapping populations which use the same or related parental lines (Table 2) [42, 43, 54]. L53 (sensitive), Al237 and C100-6 (tolerant) were the lines used in expression studies.

Both *ZmMATE1* and *ZmMATE2* localize to the plasma membrane, but gene expression and other properties differ. *ZmMATE1* shares significant identity to sorghum *SbMATE* and *Arabidopsis AtMATE* and is expressed mainly in root tips. *ZmMATE1* was the most upregulated gene in roots tips of an Al tolerant maize line under Al stress in expression studies. *ZmMATE1* was constitutively expressed in both genotypes, but was 5-10x higher in the tolerant lines. It was also strongly upregulated by Al in tolerant lines, and less so in the sensitive one. *ZmMATE1* can mediate citrate efflux. Expression levels and timing correlated with Al activate citrate release and Al tolerance. *ZmMATE1* can also increase root citrate exudation and Al tolerance in transgenic *Arabidopsis* plants [32].

ZmMATE2 shares no identify with any MATE proteins associated with Al tolerance. It also does not respond to Al and is expressed in both root tips and shoots and does not transport citrate. If and what the role of *ZmMATE2* is in Al tolerance is unknown, but it may have a role in mediating anion efflux, due to its functional ability to mediate a significant inward current. *ZmMATE2* also has and identical coding sequences in the parents of the mapping population (Al237 and L53), but several studies have shown regulation of these transporters tends to be from transposons upstream of the coding region (Table 1)[32].

Recently Maron et al [33] identified how copy number variation (CNV) in *ZmMATE1* contributes to Al tolerance and is associated with higher gene expression in a RIL population (Al237 x L53). More than 90% of maize genome shows some degree of CNV between lines. Higher copy number of ZmMATE1 *(*tandem triplication) is positively correlated with a higher Al tolerance level. However, this is very rare allele found in only 3 lines from an array of diverse maize inbred lines (166) and teosinte (23) lines; C100-6, Al237 and Il677a. This gene was tested for Al tolerance in the association study described in Chapter 2 and was not significant. However, given the low frequency of this allele (1 in 282) there would be no power to detect such a rare variant using this type of study [19, 33]. All of the lines containing this tandem triplication originate from landraces endemic to South American regions with highly acidic soil and have been used as important source of Al tolerance since early 80s. The local adaptation and selection for tolerance likely increased the allele frequency in these lines. Locally adapted maize or landraces from these acid soil regions may be a valuable resource for finding other genes that have been under selection for Al tolerance

REFERENCES

1. Miyasaka, S.C., N.V. Hue, and M.A. Dunn, *Aluminum - Handbook of Plant Nutrition*, in *Handbook of Plant Nutrition*, D.J.P. Allen V Barker, Editor 2007, CRC Press. p. 439-498.
2. Hede, A.R., B. Skovmand, and J. López-Cesati, *Acid Soils and Aluminum Toxicity*, in *Application of Physiology in Wheat Breeding*2001, CIMMYT: Mexico.
3. Ryan, P.R. and E. Delhaize, *Adaptations to Aluminum Toxicity*, in *Plant Stress Physiology*, S. Shabala, Editor 2012, CABI. p. 171-190.
4. Zhou, G., et al., *Biotechnological Solutions for Enhancing the Aluminium Resistance of Crop Plants*, in *Abiotic Stress in Plants- Mechanisms and Adaptations,*, P.A. Shanker, Editor 2011, InTech.
5. von Uexküll, H.R. and E. Mutert, *Global extent, development and economic impact of acid soils.* Plant and Soil, 1995. **171**: p. 1-15.
6. Foy, C.D., *Plant adaptation to acid, aluminum-toxic soils.* Communications in Soil Science and Plant Analysis 1988. **19**(7-12): p. 959-987.
7. Kochian, L.V., *Cellular Mechanisms of Aluminum Toxicity and Resistance in Plants.* Annual Review of Plant Physiology and Plant Molecular Biology, 1995. **46**(1): p. 237-260.
8. Horst, W.J., *The role of the apoplast in aluminium toxicity and resistance of higher plants: a review.* Zeitschrift fu¨r Pflanzenerna¨hrung und Bodenkunde 1995. **158**: p. 219-428.
9. Bloom, P.R., M.B. McBride, and R.M. Weaver, *Aluminum Organic Matter in Acid Soils: Buffering and Solution Aluminum Activity.* Soil Science Society of America, 1979. **43**(3): p. 488-493.
10. Kochian, L.V., O.A. Hoekenga, and M.A. Piñeros, *How Do Crop Plants Tolerate Acid Soils? Mechanisms of Aluminum Tolerance and Phosphorous Efficiency.* Annual Review of Plant Biology, 2004. **55**(1): p. 459-493.
11. Sivaguru, M. and W.J. Horst, *The Distal Part of the Transition Zone Is the Most Aluminum-Sensitive Apical Root Zone of Maize.* Plant Physiology, 1998. **116**(1): p. 155-163.
12. Jones DL, K.L., *Aluminum interaction with plasma membrane lipids and enzyme metal binding sites and its potential role in Al cytotoxicity.* FEBS Letters, 1997. **400**(1): p. 57-57.
13. Eticha, D., A. Stass , and W.J. Horst, *Cell-wall pectin and its degree of methylation in the maize root-apex: significance for genotypic differences in aluminium resistance.* Plant, Cell and Environment, 2005. **28**: p. 1410-1420.
14. Piñeros, M., et al., *Aluminum Resistance in Maize Cannot Be Solely Explained by Root Organic Acid Exudation. A Comparative Physiological Study.* Plant Pysiology, 2005. **137**(1): p. 231-241.
15. Kochian, L.V., et al., *Maize Al Tolerance*, in *Handbook of Maize: Its Biology*, J.L. Bennetzen and S.C. Hake, Editors. 2009, Springer New York. p. 367-380.

16. Famoso, A., et al., *Genetic Architecture of Aluminum Tolerance in Rice (Oryza sativa) Determined through Genome-Wide Association Analysis and QTL Mapping.* Plos Genetics, 2011. **7**(8).
17. Raman, H., et al., *Molecular characterization and mapping of ALMT1, the aluminium-tolerance gene of bread wheat (Triticum aestivum L.).* Genome, 2005. **48**(5): p. 781-91.
18. Magalhaes, J.V., j. liu, and C.T. Guimarães, *A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum.* Nature Genetics, 2007. **39**: p. 1156-1161.
19. Krill, A., et al., *Association and Linkage Analysis of Aluminum Tolerance Genes in Maize.* Plos One, 2010. **5**(4).
20. Famosos, A., et al., *Development of a novel aluminum tolerance phenotyping platform used for comparisons of cereal Al tolerance and investigations into rice Al tolerance mechanisms.* Plant Physiology, 2010. **153**: p. 5466-1568.
21. Jardim, S., *Comparative genomics of grasses tolerant to aluminum.* Genetics and molecular research, 2007. **11**(6): p. 1178-89.
22. Piñeros, M., et al., *Not all ALMT1-type transporters mediate aluminum-activated organic acid responses: the case of ZmALMT1 - an anion-selective transporter.* Plant Journal, 2008. **53**(2): p. 352-67.
23. Ligaba, A., et al., *Maize ZmALMT2 is a root anion transporter that mediates constitutive root malate efflux.* Plant, Cell & Environment, 2012. **35**(7).
24. Sasaki, T., et al., *A wheat gene encoding an aluminum-activated malate transporter.* Plant Journal, 2004. **37**: p. 645-653.
25. Delhaize, E., J.F. Ma, and P.R. Ryan, *Transcriptional regulation of aluminium tolerance genes.* Trends in Plant Sci, 2012. **17**(6): p. 341-8.
26. Delhaize, E., et al., *Engineering high-level aluminum tolerance in barley with the ALMT1 gene.* PNAS, 2004. **101**(42): p. 15249-15254.
27. Hoekenga, O., et al., *AtALMT1, which encodes a malate transporter, is identified as one of several genes critical for aluminum tolerance in Arabidopsis.* PNAS, 2006. **103**(25): p. 9738-43.
28. Collins, N., et al., *An ALMT Gene Cluster Controlling Aluminum Tolerance at the Alt4 Locus of Rye (Secale cereale L).* Genetics, 2008. **179**: p. 669-682.
29. Kovermann, P., et al., *The Arabidopsis vacuolar malate channel is a member of the ALMT family.* Plant Journal, 2007. **52**(6).
30. Sasaki, T., et al., *Closing Plant Stomata Requires a Homolog of an Aluminum-Activated Malate Transporter.* Plant Cell Physiology, 2010. **51**(3): p. 354-365.
31. Ligaba, A., et al., *The BnALMT1 and BnALMT2 genes from rape encode aluminum-activated malate transporters that enhance the aluminum resistance of plant cells.* Plant Physiology, 2006. **142**(3): p. 1294 -303.
32. Maron, L.G., et al., *Two functionally distinct members of the MATE (multi-drug and toxic compound extrusion) family of transporters potentially underlie two major aluminum tolerance QTLs in maize.* The Plant Journal, 2009. **61**(5): p. 728-740.
33. Maron, L., et al., *Aluminum tolerance in maize is associated with higher MATE1 gene copy number.* PNAS, 2013. **110**(13): p. 5241-5246.

34. Fujii, M., et al., *Acquisition of aluminium tolerance by modification of a single gene in barley.* Nature Communication, 2012. **6**(3): p. 713.
35. Liu, j., et al., *Aluminum-activated citrate and malate transporters from the MATE and ALMT families function independently to confer Arabidopsis aluminum tolerance.* Plant Journal, 2008. **57**(3): p. 389-99.
36. Ryan, P.R., et al., *A Second Mechanism for Aluminum Resistance in Wheat Relies on the Constitutive Efflux of Citrate from Roots.* Plant Pysiology, 2009. **149**(1): p. 340-351.
37. Mattielloa, L., et al., *Transcriptional profile of maize roots under acid soil growth.* BMC Plant Biology 2010. **10**(196).
38. PB, L., et al., *ALS3 encodes a phloem-localized ABC transporter-like protein that is required for aluminum tolerance in Arabidopsis.* Plant Journal, 2005. **41**(3): p. 353-363.
39. Prioli, A.J., et al., *Genetics analysis of aluminum tolerance in maize.* Crop Breeding and Applied Biotechnology, 2002. **2**: p. 30-33.
40. Jorge, R.A. and P. Arruda, *Aluminum induced organic acids exudation by roots of an aluminum-tolernat tropical maize.* Phytochemistry, 1997. **45**(4): p. 675-681.
41. Magnavaca, R., C. Gardner, and R. Clark, *Evaluation of inbred maize lines for aluminum tolerance in nutrient solution.* Genetic Aspects of Plant Mineral Nutrition, 1987. **27**: p. 225-265.
42. Ninamango-Cárdenas, F.E., et al., *Mapping QTLs for aluminum tolerance in maize.* Euphytica 2003. **130**(2): p. 223-232.
43. Sibov, S.T., et al., *Two genes control aluminum tolerance in maize: Genetic and molecular mapping analyses.* Genome, 1999. **42**: p. 475-482.
44. Mason, P., *Molecular and genetic investigations of aluminum tolerance in wheat and maize.* , 2005, Cornell University: Ithaca NY:. p. 141 p.
45. Rhue, R.D., et al., *Genetic Control of Aluminum Tolerance in Corn.* Crop Science, 1978. **18**: p. 1063-1067.
46. Moon, D.H., et al., *Somaclonal-variation-induced aluminum-sensitive mutant from an aluminum-inbred maize tolerant line.* Plant Cell Reports, 1997. **16**(10): p. 686-691.
47. Torres, G., et al., *A search for RFLP markers to identify genes for aluminum tolerance in maize.* Brazilian Journal of genetics, 1997. **20**(3): p. 459-465.
48. Conceição, L.D.H.C.S., C. Tessele, and J.F. Barbosa Neto, *Diallele analysis and mapping of aluminum tolerance in corn inbred lines.* Maydica, 2009. **54**: p. 55-61.
49. Magnavaca, R., C. Gardner, and R. Clark, *Inheritance of aluminum tolerance in maize*, in *Genetic Aspects of Plant Mineral Nutrition*, W.H. Gabelman and B.C. Loughman, Editors. 1987.
50. Boni, T.A., et al., *Inheritance of aluminum tolerance in maize.* Crop Breeding and Applied Biotechnology, 2009. **9**: p. 147-153.
51. Maron, L.G., et al., *Transcriptional profiling of aluminum toxicity and tolerance responses in maize roots.* New Phytologist, 2008. **279**(1): p. 116-128.

52. JL, Y., et al., *Cell wall polysaccharides are specifically involved in the exclusion of aluminum from the rice root apex.* Plant Physiology, 2008. **146**(2): p. 602-11.
53. Yang, J.L., et al., *Genotypic differences in Al resistance and the role of cell-wall pectin in Al exclusion from the root apex in Fagopyrum tataricum.* Annals of Botany, 2011. **107**(3): p. 371-378.
54. Guimarães, C.T., et al., *QTL and selection Mapping for Al tolerane in Tropical Maize*, 2009.

CHAPTER 2

ASSOCIATION AND LINKAGE ANALYSIS OF ALUMINUM TOLERANCE GENES IN MAIZE

Allison M. Krill[1,2] , Matias Kirst[3,4] , Leon V. Kochian[1,5] , Edward S. Buckler[1,2,3] , Owen A. Hoekenga[1,2]

1. United States Department of Agriculture-Agricultural Research Service, Robert W. Holley Center for Agriculture and Health, Ithaca, New York, United States of America,
2. Cornell University, Department of Plant Breeding and Genetics, Ithaca, New York, United States of America,
3. Institute for Genomic Diversity, Cornell University, Ithaca, New York, United States of America,
4. University of Florida, School of Forest Resources and Conservation, Gainesville, Florida, United States of America,
5. Cornell University, Department of Plant Biology, Ithaca, New York, United States of America

Published in PlosOne April 2010, Volume 5, issue 4, e9958

Introduction

Aluminum (Al) toxicity from acidic soil is a major constraint to worldwide crop production. Al, one of the most abundant elements in the soil, is solubilized as Al^{3+} under acidic soil conditions. This form of Al is highly toxic to plant roots. Approximately 30% of the worlds ice-free soils are acidic, 17% of which are considered arable [1]. Maize has become one of the most important grain crops grown on acidic soils due to its demand as a food crop and its ability to tolerate Al [1]. Up to a 70% reduction in maize yields have been seen in these regions due to Al toxicity [1–3]. Acid precipitation and intensive agricultural practices such as overuse of ammonia fertilizers accelerate the natural process of soil acidification, especially in the tropical and subtropical regions [4]. Soil amelioration with compounds such as lime can be used to temporarily neutralize the topsoil. However, this is not a feasible option for resource poor farmers or for subsoil acidity, and is not an economically or agronomically sustainable solution. Investing in the production of Al tolerant maize varieties and alternative management practices can contribute greatly to increased yield and sustainable crop production from acidic soils [5,6]. Therefore, an understanding of the genetic and molecular mechanisms underlying Al tolerance in maize is essential to accelerate the development of Al tolerant varieties.

The toxic effects of acid soil result from an interaction between pH and elements in the soil. Several metals, including Al and Mn, become soluble at and below pH 5.5, which causes stress in the plant. In a neutral or basic environment, Al is

found in insoluble divalent and monovalent forms of Al-oxides or Al-hydroxides, but the soluble trivalent Al^{3+} ion becomes the dominant species in an acidic environment [4]. Al^{3+} disrupts many physiological processes in plants through both apoplastic and symplastic interactions, but exact mechanisms remain elusive [7,8]. The root apex is the most sensitive part of the plant to Al because it is the site of cell division and expansion for the root [9,10]. Al-induced inhibition of root growth is the primary symptom of Al toxicity [7,9]. Reduction in root growth and function leads to increased susceptibility to other stresses, primarily drought and mineral deficiencies, due to the limited capacity of Al-intoxicated roots to acquire sufficient water and nutrition from the soil. There have been numerous mechanisms proposed for Al toxicity, but it is likely from the disruption of a number of different processes. One important site of Al^{3+} intoxication is the cell wall of the root apex [11]. In response to Al^{3+} exposure, callose formation seals off the cell walls, increasing rigidity, decreasing extensibility and preventing further transport into the cell [12]. Al^{3+} displaces Mg^{2+} and Ca^{2+}, which are required for ATPases, cell signaling, and altering or inactivating the function of many proteins [9,13]. Other possible mechanisms of Al toxicity include interference with the cytoskeleton, promotion of lipid peroxidation and blocking of Ca2+ channels [14,15]. Specifically for maize grown in the field, Al intoxication causes several stress related physiological effects, including stunting, reduced number of ears per plant, delayed flowering, and reduced biomass and total yield [3,5].

Plants have developed several mechanisms for dealing with Al toxicity, which can be classified as either external or internal tolerance mechanisms [13]. External mechanisms include differential binding of Al to the cell wall, selective permeability of the plasma membrane, formation of a plant induced pH barrier in the rhizosphere, and root exudation of chelating compounds, such as organic acids (OA) or phenolic compounds. Internal mechanisms include chelation of Al in the cytosol, compartmentalization in the vacuole, Al-binding proteins, Al tolerant enzyme isoforms, and elevated enzyme activity [4,13]. Most Al tolerance research has focused on Al induced root exudation of OA to chelate Al in the rhizosphere, where non-toxic complexes can be formed between Al and an OA such as citrate. Root exudation of OAs is a widespread response to Al in both monocots and dicots [16,17]. This mechanism has been shown to play a role in Al tolerance in several species though the activation of anion transporters in the plasma membrane [16–20].

Maize has considerable genetic variation in levels of Al tolerance, but clear physiological bases and molecular mechanisms for this tolerance remain elusive. Physiological studies found that OA exudation contributes to maize Al tolerance, but is not the only mechanism, as some Al sensitive varieties have been shown to exude high amounts of OA from the roots [21]. Differences in cell wall pectin content and degree of methylation have also been suggested to contribute to Al tolerance in maize [22]. Understanding the mechanisms of Al tolerance can accelerate the efforts to identify and incorporate superior genes and alleles into maize breeding programs. Recurrent selection has been used to develop Al tolerant maize populations with yields as much as 200% greater than susceptible lines [3,23]. However, a strong genetic by

environment (GxE) interaction and relatively low heritability of Al tolerance in maize complicates selection and has made substantial progress difficult [3,5,24].

The genetic variation for Al tolerance in maize indicates it is a complex trait, involving many genes and physiological processes [3,24,25]. Several QTL studies examined Al tolerance in maize, and suggest that about 6 loci account for ,60% of the variation in tolerance levels [21,26–28]. However, QTL in these different biparental populations are not shared, suggesting genetic heterogeneity [29]. This is not unreasonable, as the first two populations were constructed from South American maize varieties and the latter from North American lines. Transgressive segregation is seen in these three biparental mapping populations indicative of additive and/or interaction effects among alleles contributed by the two parents. Al stress was likely a powerful selective force during maize domestication and early improvement, as maize exhibits regional adaption to various levels of Al toxicity [4].

Biparental crosses used in linkage mapping, in which one or a few loci controlling Al tolerance may segregate, provide limited insight into the analysis of complex traits in general [29]. Linkage mapping has strong statistical power and is useful for understanding how and to what extent allelic effects are dependent on one another, but provides low genetic resolution unless the population is very large [30]. Alternatively, association mapping is a method for high-resolution mapping of QTL based on linkage disequilibrium (LD), and is useful for dissecting complex traits controlled by multiple QTL in species where LD decays rapidly [29,30]. Unlike linkage mapping, where only two alleles are evaluated, association mapping evaluates a greater number of alleles in a broader population. Linkage mapping uses shared inheritance of polymorphism and linked markers within families of known ancestry. Association mapping takes advantage of the historic recombination of several hundred lines, to identify common genes contributing to the trait of interest. The LD structure of the gene is essential in association mapping. This approach allows evaluation of genes from smaller sampled regions, within the range of LD decay, instead of requiring complete candidate gene sequencing. This method requires three data types: phenotypic trait information, genotypic data from or near the gene of interest and an understanding of population structure within the test panel. Beyond the requirement for prior molecular knowledge, the other principal disadvantage of association mapping is that spurious marker trait associations can arise from population structure. However, we can identify many of these false positive results via a mixed linear model (MLM) approach, which takes population structure and varietal relatedness into account [31]. The combination of association mapping and linkage mapping can provide both the power and resolution needed for detecting QTL of interest.

In this study, we used an integrated approach combining association mapping with linkage mapping to identify and evaluate candidate Al tolerance genes in maize. Without positively identified mechanisms or biochemical pathways involved in Al tolerance, selection of candidate genes requires knowledge based on previous studies and proposed mechanisms. We tested 21 candidate genes for association with Al

tolerance, in a maize diversity panel of 282 inbred lines, using the MLM approach discussed earlier [31–34]. Candidate genes were screened in a subset of 27 diverse lines (DL), selected to be representative of the genetic and phenotypic diversity in the association panel, in order to identify highly polymorphic regions for further association studies. Due to strong GxE effects in field studies of Al tolerance, selection or testing of tolerance in pots of acid soil or hydroponics solutions is a quick and efficient way to determine tolerant and sensitive lines in maize while controlling for environmental effects [35]. Al tolerance levels were measured, as net root growth (NRG) in nutrient solution containing a toxic level of Al [36]. Several genes were found to be associated with NRG under Al stress and subsequently confirmed using linkage analysis.

Results

Phenotypic data

Phenotypic data for Al tolerance in the maize association panel was collected as net root growth (NRG) in a hydroponic nutrient solution with or without a toxic level of Al^{3+} [37]. Al stress measurements were taken before and after 2 days of stress in a hydroponic solution containing (27 uM Al^{3+}) at pH 4.0. A control treatment was carried out over the same time period in an identical hydroponic solution, containing no Al3+. A wide range of tolerance levels is seen in this panel for both control NRG and Al treated NRG (Figure 1 and Table S1). Mean NRG under control treatment and under Al stress was 50.04 mm and 37.79 mm, respectively. Differences between the two groups were highly significant (p= 1.56). Mean correlation between replications was 42.5% in Al stress treatments and 37.3% in control treatment replications. Narrow sense heritability (h^2) for NRG in the Diversity Panel, calculated using the relatedness (K) matrix, was between 30 and 32% in the Al stress environment and 22% without Al stress. Broad sense heritability (H^2) for NRG under Al stress and without Al stress was 41% and 37%, respectively. These heritability estimates for net root growth in seedlings are similar to those observed in maize breeding programs for enhanced tolerance to low pH [5].

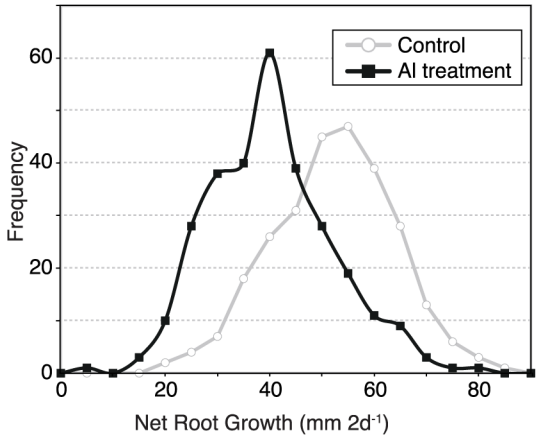

Figure 1. Distribution of Net Root Growth (NRG) in the Maize Association Panel.

Bins reflect grouping the inbred lines by 5 mm 2 d21 increments of root growth under both Al stress (circles) and control (squares) treatments. Reported values are Least Squares Means calculated from five replicate experiments for Al stress treatment or three replicate experiments for control treatment.

Genotypic data

Genotypic data for association mapping came from polymorphisms identified in candidate gene sequences. Genes were chosen based one of two factors: as responsive to Al-stress treatments according to gene expression analysis or by sequence similarity to Al tolerance genes found in other species (Table 1). Genes throughout the remainder of this study are referred to by the Gene ID listed in Table 1. Thirteen candidate genes were identified as differentially responsive to Al stress treatments in root tips, from Al tolerant and Al sensitive maize lines, in previous studies [38]. Eight candidate genes were chosen by comparative genomics based on their contribution to Al tolerance in related grass species. *TaALMT1* (Aluminum activated malate transporter) is the major Al tolerance gene in wheat (Triticum aestivum) and is the first true Al tolerance gene identified in any plant [39]. Seven maize genes homologous to *TaALMT1* were examined and are referred to as *ZmALMTx*. One gene homologous to *AltSB*, the major Al tolerance gene in Sorghum bicolor, is referred to as *ZmASL* (Zea mays *AltSB* – like) [20]. The selection of genes using a comparative genomics approach is based on evidence suggesting many agronomically important traits, such as Al tolerance, may be controlled by orthologous loci in related grasses or more distant species [40]. For example, genes related to *TaALMT1* from wheat have been demonstrated as Al tolerance genes in Arabidopsis and rye [18,39,41,42], while genes related to *AltSB* from sorghum have been demonstrated as Al tolerance genes in Arabidopsis and barley [20,40,42,43].

Gene ID	Gene Name	MAGI ref seq#	Length (bp)	Lines (#)	Sites (#)	Chr (#)	ctg (#)	BIN (#)
ME	Malic Enzyme	3.1_47445	626	255	12	6	282	6.05
FE	Iron-responsive transporter-like	3.1_61976	503	240	11	2	106	2.08
ANTI	Major facilitator superfamily antiporter	3.1_69188	549	246	14	6	285	6.05
ABC	ABC transporter-like protein	3.1_80359	373	251	21	2	106	2.08
ISL	Isocitrate Lyase	3.1_108586	526	228	6	7	322	7.03
AUX1	Amino acid permease AUX1	3.1_112316	526	206	23	2	106	2.08
SAHH	SAH hydrolase	4.0_116767	500	206	6	4	160	4.03
P450	Cytochrome P450	4.0_145633	786	246	14	3	138	3.06
PME	Pectin methylesterase	4.0_158804	454	270	3	1	49	1.08
PI3K	Phosphatidylinositol 3-kinase	4.0_112182	590	254	10	4	172	4.05
OO2	Germin2 (oxalate oxidase)	4.0_67335	729	182	26	10	399	10.03
IDH	Isocitrate dehydrogenase	4.0_48631	1061	253	14	4	173	4.05
FUM	Fumerase	4.0_35824	646	269	8	1	14	1.04
AL1	ZmALMT1	3.1_92675	476	182	31	10	412	10.04
AL2	ZmALMT2	3.1_93496	487	199	17	10	412	10.04
AL3	ZmALMT3	3.1_92049	633	193	9	5	252	5.07
AL5	ZmALMT5	3.1_811363	743	215	24	10	412	10.044
AL8	ZmALMT8	3.1_90876	504	208	26	5	247	5.06
AL9	ZmALMT9	3.1_6591	288	263	2	5	214	5.03
AL16	ZmALMT16	3.1_36360	HAP*	285	1	10	415	10.06
ASL	ZmASL (region 1)	3.1_41691	1179	278	32	1	9	1.02
ASL	ZmASL (region 2)	3.1_41691	715	240	21	1	9	1.02

Thirteen genes were selected from gene expression analysis, while another eight came from comparative genomics. Genes were identified from genome survey sequence contigs created by the MAGI Project. MAGI build version and reference number are reported. ''Length'' describes the total length of sequence used for polymorphic site identification. ''# Lines'' refers to the number of entries with sufficient information to include in the association analysis. ''# Sites'' refers to polymorphisms that occurred at greater than 10% frequency. Physical-genetic map locations for each candidate gene are reported according to chromosome, genomic sequencing contig and genetic map bin. Genetic bins that appear in bold represent those under previously reported Al tolerance QTL. Gene AL16 was evaluated by a large indel (*HAP) rather than by gene sequence. Gene ASL underwent two rounds of sequence analysis.
doi:10.1371/journal.pone.0009958.t001

Table 1. Candidate Al tolerance genes evaluated by association mapping.

Information regarding the genes used in association mapping is shown in Table 1. A region of high polymorphism in each gene (based on preliminary sequencing in the 27 DL) was sequenced in the association panel. Polymorphisms with frequency <10% were extracted from the sequences for analysis, giving a total of 331 sites across all genes and an average of 15 sites per gene (Table S2). Reference sequence, length of sequence, number of lines with sufficient quality sequence and the physical map location of each gene are shown [44,45]. Given the heterogeneity in the rates of LD decay for these genes, size of the genes, and the possibility for distant regulatory elements, these polymorphism surveys are not intended to be comprehensive surveys of polymorphism. Instead, the sequencing results presented here are a representative sample that enables us to efficiently screen a large number of loci and identify markers with strong associations to Al stress tolerance.

Association mapping

The mixed linear model (MLM) was used for association mapping [31]. The MLM accounts for multiple levels of relatedness, defined as population structure (Q) and a pairwise kinship matrix (K), to control for both Type I and Type II errors [31]. A General Linear Model (GLM) including Q was also tested. Both models, GLM and MLM, were applied to NRG under Al stress and NRG under no Al stress. NRG under

no Al stress was also used as a fixed effect covariate in the MLM model, Q+K+C (Table 2 & Table S3). This model was used to evaluate relative root growth, which is frequently used as a measurement of Al tolerance.

Candidate Genes	# Sites	GLM (Q)	MLM (QK)	MLM (QKC)	Max r^2 (Model)
ME	12	0	0	3	1.3% (QKC)
ISL	6	0	0	0	1.2% (QKC)
SAHH	6	0	0	2	2.1% (QKC)
AL2	17	1	2	0	2.7% (QK)
ASL	53	26	10	5	2.0% (QK)
PME	3	0	1	1	1.6% (QK)
FDR $p<0.01$		40%	24%	34%	
h^2		n/a	0.30	0.32	
H^2			0.41		

GLM and MLM analyses were used to evaluate the 21 candidate Al tolerance genes, using the net root growth trait collected from Al treated plants. These models incorporated the population structure (Q) of the Diversity Panel, the relative kinship (K) of the Diversity Panel and net root growth of the Diversity Panel grown without Al stress as a fixed effect covariate (C). The GLM model used only factor Q, while the MLM models used factors Q+K and Q+K+C. Six candidate genes gave significant results and are shown, with the number of significant sites ($p<0.01$) identified per locus for each model. The maximum value for variance explained by a marker within a gene in any model is reported. False Discovery Rates were empirically calculated for each model based on 1,095 random SNPs throughout the genome and are expressed as percentages. Narrow (h^2) and broad sense (H^2) heritability estimates were generated for each trait based upon variance estimates from the MLM.
doi:10.1371/journal.pone.0009958.t002

Table 2. Evaluation of association mapping results by ANOVA.

Six genes had statistically significant associations ($p<0.01$) with NRG under Al stress and were selected for further study in F_2 linkage populations: *malic enzyme (ME); isocitrate lyase (ISL); SAH hydrolase (SAHH); ZmALMT2 (ALMT2); ZmASL (ASL); pectin methylesterase (PME*) (Table 2). Complete results from the MLM analysis can be found in Table S3. ISL was statistically significant at a less stringent value (p,0.05) for Al stress. In order to estimate the number of expected false positives due to multiple testing of sites, a false discovery rate (FDR) was calculated for each model using 1,095 random SNPs throughout the genome. FDR allows for the comparison of significant sites in our candidate genes to those we would expect to see by random chance alone. Based on the FDR values for the MLMs, about 24% of the sites under the Q+K model and 34% under the Q+K+C model under Al stress could be accounted for by false positives. Given this high rate for false discovery, it is crucial to test the connection between the six genes with putative association to Al stress tolerance using an independent line of reasoning.

Linkage mapping

If the association analysis truly identified Al tolerance genes, then the

associated SNPs should explain significant variance for Al tolerance in segregating populations. Linkage mapping could therefore be used to test the results of association mapping. Linkage to Al tolerance was tested for the six genes listed in Table 2 using three F2 populations. F2 populations were phenotyped in the same manner as the association panel and genotyped for the sites of interest (Table S4). These F2 populations were constructed so that each would segregate for polymorphisms associated with two putative Al tolerance loci: *ZmASL* and *SAHH* within B73xCML247; *ME* and *ISL* within B73xCML333; *PME* and *ZmALMT2* within B73xNC350 (Figure 2). A comparison of means for each allelic class suggested that the polymorphisms tested at *ZmASL*, *SAHH*, *ME*, and *ZmALMT2* were significantly associated with Al tolerance (Figure 2). However, allelic means for *ISL* and *PME* were equivalent no matter the state, suggesting that the polymorphisms tested were not associated with Al tolerance. Linkage was tested by GLM for the 4 putative Al tolerance genes, assuming complete dominance (*ZmASL*, *SAHH* and *ZmALMT2*) or additive gene action (*ME*; Table 3). These results indicate that small effect (3–6% variance explained) QTL exist for Al tolerance at these four loci. No significant interactions between Al tolerance genes were found, suggesting that epistasis is not at work. The identification of *ISL* and *PME* as Al tolerance gene based on association mapping were likely false positives, as there was no linkage to Al tolerance differences with the polymorphisms tested in F2 populations, and is consistent with our expectations based on the FDR calculation.

The four genes with significant association and linkage to Al tolerance, Zea mays *AltSB*-like (*ZmASL*), *S-adenosyl-L-homocysteinase* (*SAHH*), *Malic Enzyme* (*ME*), and *ZmALMT2* (*ALMT2*), are described in further detail below. All genes possessed more than one statistically significant polymorphism associated with Al tolerance differences. The complete coding sequences for the *ZmASL*, *SAHH* and *ME* genes were characterized in the 27 DL to look for other regions of interest such as non synonymous sites, alternative splicing, and protein structure modifications. *ALMT2* was not sequenced in the 27 DL subset due to constraints caused by abundant paralogs within the *ZmALMT* family. Individual sites in these genes explain only about 2% of the phenotypic variance in the association panel, but confer 13%–20% increase in NRG.

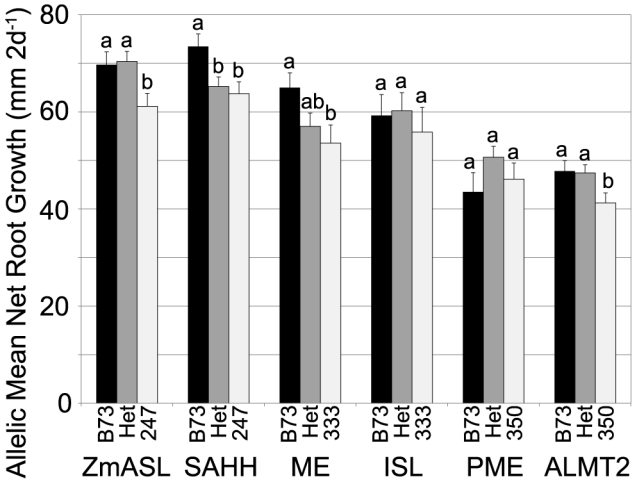

Figure 2. Linkage mapping validation of candidate Al tolerance genes. Six candidate Al tolerance genes were evaluated using three F2 linkage populations: B73xCML247, B73xCML333, and B73xNC350. Mean Net Root Growth values for each allelic state are reported, abbreviated as B73 for the B73 homozygous class, Het for the heterozygous class, and the numerical portion of the non-B73 parent name for the other homozygous class; error bars reflect standard error. Student's t-test was used to evaluate differences between allelic classes within each F_2; differences significant at $p<0.05$ are indicated with letter codes.

Factor	**DF**	**SS**	**F**	**P**
ALMT2 (dom)	1	1241.43	6.73	0.0106
Error	125	19482.91		
Model Total	126	20531.35	Adjusted r^2 =	0.043
Factor	**DF**	**SS**	**F**	**P**
ME (add)	2	2721.62	3.29	0.0403
Error	127	52484.35		
Model Total	129	55205.97	Adjusted r^2 =	0.034
Factor	**DF**	**SS**	**F**	**P**
SAHH (dom)	1	1903.55	8.58	0.0040
Error	129	28631.73		
Model Total	130	30535.28	Adjusted r^2 =	0.055
Factor	**DF**	**SS**	**F**	**P**
ZmASL (dom)	1	2037.51	8.44	0.0044
Error	124	29951.96		
Model Total	125	31989.47	Adjusted r^2 =	0.056
Factor	**DF**	**SS**	**F**	**P**
SAHH (dom)	1	1445.72	6.94	0.0097
ZmASL (dom)	1	1586.91	7.62	0.0068
Error	106	22085.66		
Model Total	108	25684.81	8.57	0.0004
			Adjusted r^2 =	0.123

GLM analysis was used to evaluate whether SNP markers within candidate Al tolerance genes explained significance variance for Al tolerance observed in F_2 populations. Gene action was modeled as either additive ("add") or dominant ("dom") based on allelic means. The variance explained by each significant SNP is reported. As both *SAHH* and *ZmASL* were significantly associated with Al tolerance for the B73×CML247 population, a summary model is reported. DF: Degrees of Freedom; SS: Sum of Squares; F: F ratio; P: P value.

doi:10.1371/journal.pone.0009958.t003

Table 3. Evaluation of linkage mapping results by ANOVA. GLM analysis was used

to evaluate whether SNP markers within candidate Al tolerance genes explained significance variance for Al tolerance observed in F2 populations. Gene action was modeled as either additive (‘‘add’’) or dominant (‘‘dom’’) based on allelic means. The variance explained by each significant SNP is reported. As both *SAHH* and *ZmASL* were significantly associated with Al tolerance for the B73xCML247 population, a summary model is reported. DF: Degrees of Freedom; SS: Sum of Squares; F: F ratio; P: P value.

Al tolerance gene: ZmASL

ZmASL, which is highly similar to the Al-activated citrate transporter from sorghum, is described in Figure 3. Figure 3A shows the gene organization for ZmASL, including exons, introns and nonsynonymous sites, based on the genomic sequence of the 27 DL. Total length of ZmASL sequenced in the 27 DL was about 6 kb, including 11 exons, both 59 and 39 UTRs and an upstream region containing a 300 bp MITE insertion. The common polymorphisms (frequencies ≥10%), which are responsible for 12 amino acid substitutions, are shown. The 43 rare amino acid substitutions, insertions or deletions (<10% frequency) are not shown. Many of the rare polymorphisms are found in only one of the 27 DL (CML247 was responsible for 21 sites).

Two minimally overlapping regions of this gene were sequenced in the association panel, covering a total of 1.7 kb (Figure 3B). These regions represent the first three exons and part of the fourth, and were selected, as they were highly polymorphic for both synonymous and nonsynonymous sites, including 7 of the 12 common amino acid substitutions. Based on this sequence the remaining 5 amino acid substitutions were inferred from haplotype structure. The MITE insertion in the 59 UTR was also scored in the panel. Altered gene expression in the AltSB gene is associated with the number of MITE insertions in the regulatory region of that gene [20]. However, the MITE found in ZmASL was not associated with NRG.

We detected 11 sites that were significantly associated with NRG under Al stress in the MLM models, as shown in Figure 3B. All of the significant sites occur in introns. A 120 bp indel (site #47) in the second intron showed the highest statistical significance and was in high LD with several of the other significant sites. A total of three independent sites (R^2=0.2) were significantly associated with NRG in the region sequenced (Figure 3C). Each significant site in the Al stress statistical models explains between 1.5% and 2.7% of the total phenotypic variance observed in the association panel. However, the most significant site has an effect estimate that increases NRG 16% over the two days of Al stress. The 120 bp indel (site #47) was used for the linkage analysis in the B73xCML247 F_2 population, where it was correlated with a 15% increase in NRG. The superior allele found in B73 appeared to be fully dominant to the inferior allele found in CML247 (Figure 2).

Figure 3D shows the predicted transmembrane protein structure of this gene containing 10 putative transmembrane domains [46]. The approximate locations of the

12 common amino acid substitutions on the protein are shown. However, none of these polymorphic sites were significantly associated with NRG under Al stress.

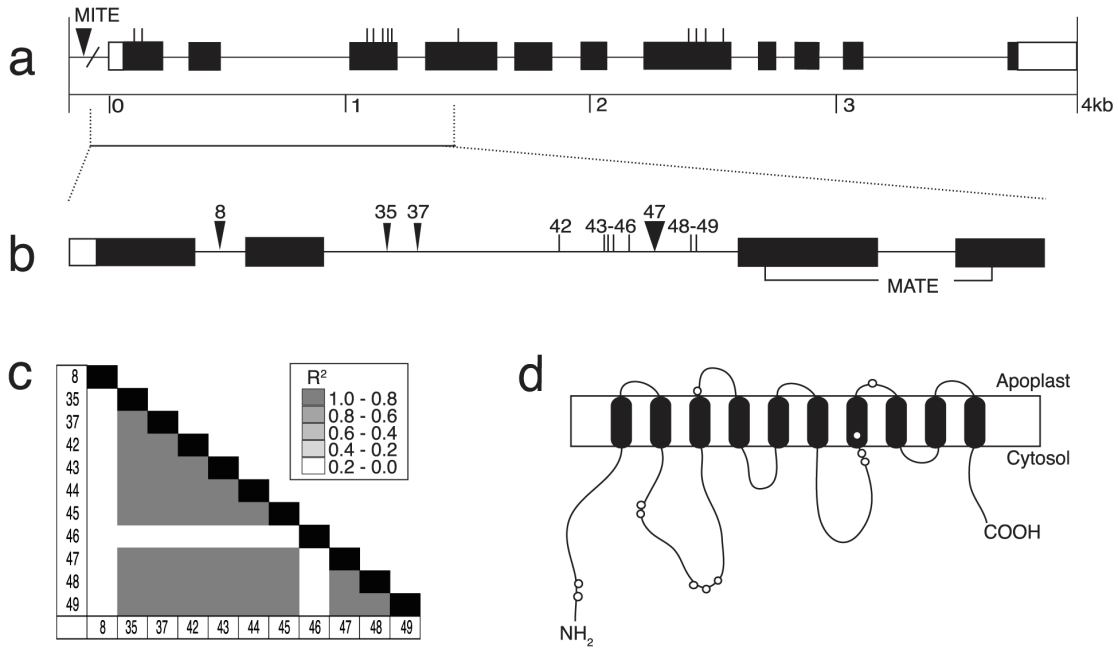

Figure 3. Characterization of *ZmASL*.
(a) Predicted gene structure for the *ZmASL* locus is shown, with exons as black boxes, introns as thin lines, and UTRs as open boxes. The approximate location of 12 amino acid substitutions or additions that occur at greater than 10% frequency among alleles are shown with vertical lines above the exons, based on complete *ZmASL* sequencing performed in the 27 DL subset. (b) A focus region of *ZmASL* was sequenced in the association panel. The polymorphisms that were identified as significant by the association analyses are shown – SNPs as vertical lines, indels as triangles – and are referred to by number. The conserved MATE domain is highlighted in exons 3 and 4. (c) Linkage disequilibrium plot for the eleven significant polymorphisms. High linkage disequilibrium exists between nine of the eleven associated polymorphisms. (d) An estimate for the transmembrane structure of *ZmASL*, where open circles indicate the approximate locations for the 12 amino acid substitution/insertions detected within the gene

Al tolerance gene: SAHH

The complete predicted coding sequence for SAHH was sequenced in the 27 DL. This 2.5 kb region includes three exons and the 39 UTR (Figure 4A). Only one amino acid substitution was observed in the 27 DL gene sequences and is encoded by a triallelic SNP (#5). The region sequenced in the association panel spanned most of the first exon, including this amino acid substitution (Figure 4B). We observed 6 SNPs and no indels in this portion of the first exon. Two nonsynonymous SNPs (#1 and #2), in high LD (R2≥0.8), were significant for NRG under the Q+K+C model (Figures 4B & 4C). The triallelic SNP (#5) was significant at the p,0.05 level, and leads to either a synonymous (Glu for Glu) or conservative (Asp for Glu) amino acid substitution. The

triallelic SNP was in moderate LD with associated SNP #1 and in little or no LD with SNP #2. The two highly significant sites (#1/#2 and #5) explain between 1.8 and 2.1% of the phenotypic variation and confer up to a 13% increase in NRG under Al stress.

Instead of utilizing one of the associated SNPs, an indel polymorphism identified in the first intron during whole gene sequencing was used for the linkage analysis. The choice of the indel provided us a simple PCR based assay for genotyping and took advantage of the difference in genetic resolution between association mapping and linkage mapping. Far fewer recombination events were captured in the F2 population than in the association panel, thus an indel that was not scored in the complete association panel was equally useful for linkage analysis. This indel was correlated with a 13% increase in NRG in the B73xCML247 F_2 population (Figure 2), the same relative increase as we attributed to SAHH by association mapping. The inferior allele of SAHH found in CML247 was fully dominant to the superior allele from B73 (Figure 2).

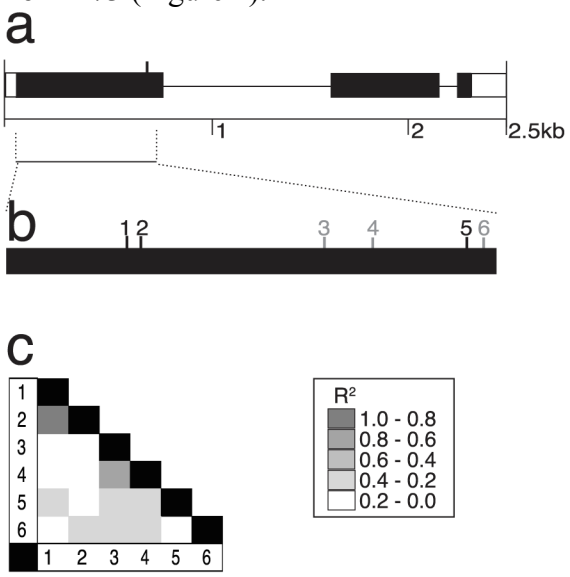

Figure 4. Characterization of *SAHH*.
(a) Predicted gene structure for the *SAHH* locus is shown, with exons as black boxes, introns as thin lines, and UTRs as open boxes. A single amino acid substitution was detected from complete gene sequencing in the 27 DL subset and is indicated by the vertical line in the first exon. (b) A focus region of *SAHH* was sequenced in the association panel. Six SNPs were detected in the association panel and are referred to by number. Polymorphisms 1, 2 and 5 were identified as significantly associated with aluminum tolerance differences and are shown in black; non-significant sites are shown in gray. Site #5 corresponds to the triallelic SNP that causes the single amino acid substitution detected. (c) Linkage disequilibrium plot for all polymorphisms detected in the focus region at *SAHH*. High linkage disequilibrium exists between sites 1 and 2, while relatively low linkage disequilibrium exists through the rest of the gene.

Al tolerance gene: ME

The complete predicted gene sequence for ME, including both 59 and 39 UTRs

and a farther 59 region with two large insertions was sequenced in the 27 DL (approximately 5 kb; Figure 5A). Like SAHH, we saw very little nucleotide diversity in the ME sequences – only one amino acid substitution was seen in more than one line, located in the last exon. Three rare amino acid substitutions were seen in one line. The 59 UTR, the first exon, and most of the first intron were sequenced in the association panel (Figure 5B). Three sites (#4, #7, and #11) were associated with NRG under Al stress, two in high LD were found in the second intron and one independent site in the first exon (Figure 5C). The most significant independent site (#4) explains 1.4% of the variance in the association panel, which translates to an 18.4% increase in NRG under Al stress. We also examined the far upstream region in the association panel, which contained two large indels, but no significant associations were found.

Site #1, an indel in the first intron, was used for the linkage analysis of the B736CML333 F2 population. The superior B73 allele of ME was correlated with a 21% increase in NRG, similar to the effect seen in the association mapping. The heterozygous class was intermediate in phenotypic effect, unlike that seen with ZmASL or SAHH, suggesting that the mode of action was additive rather than dominant.

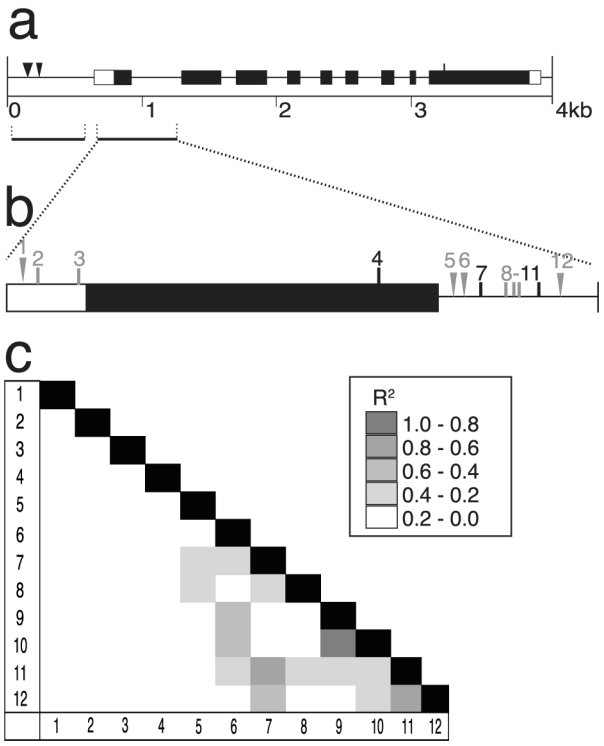

Figure 5. Characterization of ME.
(a) Predicted gene structure for the *ME* locus is shown, with exons as black boxes, introns as thin lines, and UTRs as open boxes. Two indels and an SNP were detected in sequencing the complete gene in 27 DL subset. (b) A focus region of *ME* was sequenced in the association panel. Four indels and eight SNPs were detected; three SNPs were significantly associated with aluminum tolerance and are shown in black. (c) Linkage disequilibrium plot for all polymorphisms detected within the focus region

at *ME*. No linkage disequilibrium exists within the 5' end of the focus region, while moderate linkage disequilibrium exists among several of the 3' end sites.

Al tolerance gene: ALMT2

We evaluated seven members of the *ZmALMT* gene family by association analysis. Only *ZmALMT2* (*ALMT2*) gave a significant result (Figure 6). The gene model shown in Figure 6A is based on B73 sequence information only because of sequencing constraints due to paralogs within the *ZmALMT* family. Two SNPs in this gene (#2 and #12) were associated with NRG under the Q+K model (Figure 6B). These SNPs were independent of each other (R^2=0.2), although LD was moderate to extensive between most of the SNPs found at this gene (Figure 6C). The most significant SNP explains 2.7% of the variation in the panel and confers a 20.2% increase in root growth.

Site #11, an indel that was not associated with Al tolerance, was used for the linkage analysis in the B73xNC350 F_2 population. Like *ZmASL* and *SAHH*, the superior allele found in B73 was fully dominant to the allele found in the other parent. However, unlike *ZmASL* and *SAHH*, the enhancement in NRG correlated with *ALMT2* was somewhat smaller (15%) in the linkage population than expected from the association population (20.2%).

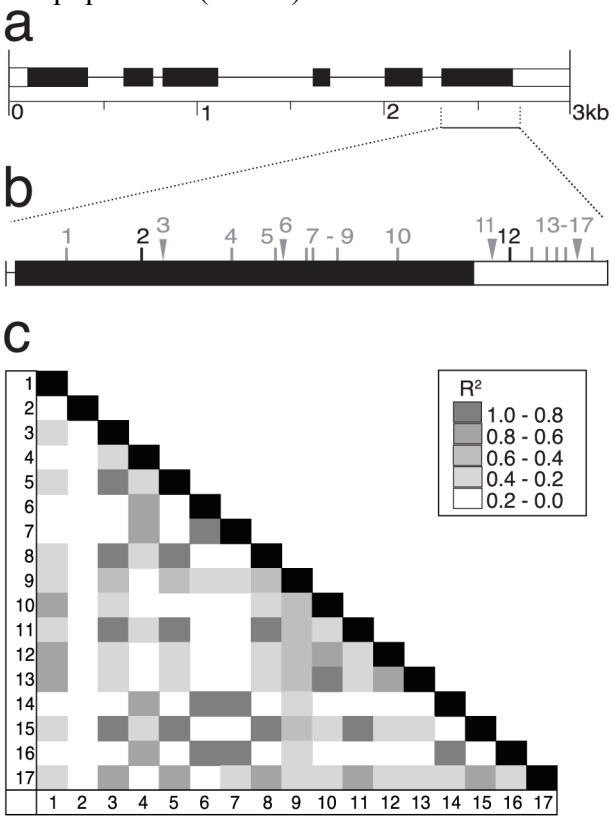

Figure 6. Characterization of *ZmALMT2*. (a) Predicted gene structure for the *ZmALMT2* locus is shown, with exons as black boxes, introns as thin lines, and UTRs as open boxes. For ease of presentation, polymorphisms detected in the 27 DL

complete gene sequencing are not shown. (b) A focus region of *ZmALMT2* was sequenced in the association panel. Four indels and thirteen SNPs were detected; two SNPs were significantly associated with aluminum tolerance and are shown in black (#2, #12). (c) Linkage disequilibrium plot for all polymorphisms detected within the focus region at *ZmALMT2*. Site 2 is associated with aluminum tolerance differences but is in linkage equilibrium with all other sites, while site 12 has moderate linkage disequilibrium with many sites within the focus region.

Discussion

We found four new genes that may contribute to Al tolerance in maize by integrating several approaches. Candidate genes were selected based on comparative genomics and gene expression analysis, which we evaluated with association and linkage mapping. Two of these genes, *ZmALMT2* and *ZmASL*, are very similar to Al activated organic transporters that play crucial roles in determining Al tolerance in other species [20,39–43]. Whether these new maize genes are also membrane transporters has yet to be determined. *ME* and *SAHH* are involved in several central metabolism reactions and speculations on their contribution to Al tolerance can be made based on previous studies [47,48]. The low heritability and complex nature of Al stress tolerance makes it challenging for both genetic improvement and genetic dissection. This complexity highlights the value for molecular markers for use in breeding programs for Al. tolerance. To maximize their efficacy, molecular markers should be closely linked to major tolerance loci, so that markers are highly concordant with the desired genotypes. Given the limited amount of DNA sequence obtained for association mapping within many of the genes we investigated, we cannot positively identify these particular polymorphisms as causative without further investigation. However, they are most likely in LD with causative sites or contribute to an allelic series controlling Al tolerance, and therefore will be highly useful as markers for selection of Al tolerance materials. We demonstrated that using polymorphisms identified during gene sequencing as markers for linkage analysis allowed us to confirm the identification for four of the six putative Al tolerance genes. We utilized sites that were both significantly associated with Al tolerance differences in the association panel and sites not significantly associated, taking advantage of the difference in genetic resolution between association and linkage mapping. It was important to use linkage mapping to test the genes identified from association mapping, as we predicted a high rate for false discovery based on empirically calculating an FDR. Each of the Al tolerance loci produced similar phenotypic effects in both the association panel and F2 populations (13–20% increases in NRG). While none of these new Al tolerance genes represent major effect QTL, combining multiple small QTL can make a significant impact to enhance the desired trait. In the B73xCML247 F_2 population, combining the elite alleles of *ZmASL* and *SAHH* enhanced net root growth by 30% (Tables 3 and S4).

ZmASL (Zea mays AltSB like) is a maize gene homologous to *AltSB*, the major Al tolerance gene from sorghum and is a member of the Multidrug And Toxic

Compound Extrusion (MATE) family of transporters [20]. Both proteins are predicted to contain 10 putative transmembrane domains. It is unknown whether the *ZmASL* gene mediates Al-activated root citrate efflux, as is the role of AltSB in sorghum. Although many sites in *ZmASL* were associated with NRG under Al stress, none were amino acid substitutions. The significant sites we detected may be in LD with regulatory elements of the gene, as is the case in *AltSB*, where polymorphisms in the promoter help to determine the level of gene expression [20]. *ZmASL* contained the most significant independent sites of any gene tested, but also contained extensive LD among many of the other significant sites. Fortunately, future experiments to evaluate the relationship of ZmASL with Al tolerance will be relatively straightforward given the presumed gene function.

SAHH, S-adenosyl-L-homocysteine hydrolase, is an enzyme that removes the feedback inhibitor of SAM (S-adenosylmethionine) mediated methylation in any organism [47]. Any enzymatic process that requires high rates of SAM-mediated methylation will also require high *SAHH* activity, including DNA/RNA modification, nucleic acid metabolism, and synthesis of cell wall constituents [49]. *SAHH* has a high degree of sequence conservation among eukaryotes [49]. In plants, *SAHH* is a cytokinin binding protein in plants, induced by auxin and cytokinin, and has been associated with salt-stress response in spinach and sugar beets and viral resistance in Arabidopsis [47,50]. The isoform of maize *SAHH* we examined was previously found to be highly expressed in root tips under Al stress [38]. The connection of *SAHH* to Al tolerance could come through any of several mechanisms due to the broad range of processes the enzyme is involved in. However, given recent reports on the correlation pectin methylation in cell walls with Al tolerance and Al exclusion, it is certainly possible the *SAHH* contributes to Al tolerance differences via cell wall modification [22,51].

NADP-ME (ME) catalyzes the conversion of malate to pyruvate. The maize ME examined in this study was the cytosolic rather than plastidic isoform of the enzyme. Maize Cyt-*ME* is highly similar to Cyt-*ME* found other in C3 and C4 plant species. This isoform was found to be expressed in the embryo and emerging roots, with expression responsive to hypoxia and drought [48]. High malate and other organic acid concentrations are optimal for activity of the cytosolic isoform and not inhibitory as is the case in plastidic isoforms of *NADP-ME* [48]. There is strong evidence that Al-activated release of malate underlies wheat Al tolerance [9,17,39]. Malate appears to chelate and detoxify Al in the apical rhizosphere or the apoplastic space. ME may help regulate malate concentration in the cytosol, which could connect to Al tolerance either through OA efflux or internal detoxification of Al via Al-OA chelation.

ME was unusual among the genes we examined as the results from linkage and association studies were opposite in direction, while still both highly significant. In the association mapping, three significant sites were identified – a site in the first exon (SNP #4), which was in linkage equilibrium with all other sites, and two sites in the

second intron (SNP #7 and SNP #11), which were in high linkage disequilibrium with all of nearby the SNPs (Figure 5c). Based on these sites, we predicted that the B73 allele would be inferior to the CML333 allele. However, in the linkage mapping B73 was superior to CML333 (Figure 2). One possible explanation is that an allelic series exists at ME that was not observed in the polymorphic sites studied in the association panel. However, an allelic series could be detected in the larger linkage blocks of the segregating population. We see evidence of allelic series in several other candidate genes studies in maize, such as su1 and LcyE, that also exhibit these inconsistencies between association and linkage mapping [52].

ALMT2 is related to transport proteins that have been found to contribute to Al tolerance in *Triticum aestivum, Arabidopsis thaliana*, and *Brassica napus*, and are either activated or show enhanced malate efflux in response to external Al3+ [18,39,53–56]. It is proposed that binding of Al3+ to the transporter induces a conformational change, opening the anion channel [55,56]. However, not all *ALMT* family proteins are Al-activated or important for Al tolerance processes. *AtALMT9* encodes a vacuolar malate transporter, instead of being localized to the plasma membrane like *AtALMT1* [18]. Unlike *AtALMT1*, *AtALMT9* is completely unresponsive to Al treatment [56]. The first *ZmALMT* family member to be characterized, *ZmALMT1*, transports inorganic anions and not malate, and is not activated by exogenous Al3+ [54]. Based on its transport properties and expression, *ZmALMT1* was determined not to be involved in maize Al tolerance. This is consistent with the results from the association analysis, as *ZmALMT1* was not associated with NRG under Al stress. Only *ZmALMT2* was found to be significant for Al tolerance of the seven *ZmALMT* genes that we evaluated by association analysis. Future work on *ZmALMT2* will include a biophysical characterization of the protein to verify that it does encode an Al activated OA transporter.

In summary, we used association mapping to evaluate twenty-one candidate Al tolerance genes. Linkage mapping was used to test six putative Al tolerance genes found from association mapping; this was especially important given the high predicted FDR for the association mapping. Linkage mapping supported four of the six genes as true Al tolerance genes. These four genes, ZmASL, ZmALMT2, ME and SAHH, are excellent candidates for future laboratory and field-based studies on Al tolerance in maize. Although the most significant polymorphisms explain less than 3% of the variation seen in the association panel, our best marker can increase NRG up to 20%. If this increased root growth transfers to field trials, integration of these markers could substantially improve maize root growth and overall maize yield under Al toxic conditions.

Materials and Methods

Germplasm

The maize association population has been previously described [31,34]. Linkage mapping experiments were conducted with three independent F_2 populations

derived from B73 and one of three other inbred lines from the maize association population (CML247, CML333, NC350). Non-B73 parents were selected on the basis of genotype information for the candidate Al tolerance genes.

Phenotypic analysis

Maize seeds were germinated in either autoclaved sand or on filter paper, moistened with deionized water, for 3–5 d at 28uC in continuous darkness. Seedlings were rinsed and placed into sample cups suspended in 8L vessels containing a nutrient solution without Al^{3+} at pH 4.0, for 1 d, for acclimation to hydroponic conditions [37]. When plants were placed into hydroponic culture, secondary roots were removed to promote measurement of primary seminal root growth only. Tubs were aerated and plant grown under controlled environmental conditions (26uC day/24uC night, 16 h/8 h photoperiod). After 24 hrs of acclimation, initial root growth (IRG) measurements were taken using rulers with millimeter precision and solutions were replaced with Magnavaca nutrient solution containing (27 μMAl^{3+}) at pH 4.0 (Al stress treatments) or Magnavaca nutrient solution containing no Al^{3+} at pH 4.0 (control treatments), for 2 d. After 2 d of Al stress final root growth (FRG) measurements were taken. Net root growth (NRG) was calculated as FRG – IRG.

Five replicate experiments were performed for Al stress treatments, while three replicates were performed for the control treatment. In each experiment, 3–4 individuals for each of the 282 inbred varieties in the association panel were phenotyped in each replicate experiment. Least squares means (LSmean) for both traits were calculated in SAS version 9.1 for Windows (SAS Institute Inc., Cary, NC, USA) and used as the phenotypic values in all models (Table S1).

F2 linkage populations were phenotyped in a similar manner, with (27 μM Al^{3+}) at pH 4.0, with the modification that 200 F_2 individuals were evaluated for each cross plus parental checks (n=10). Leaf tissue was collected for DNA extraction and genotypic analysis after FRG measurements. Measurements of NRG under control and stress treatments are found in Table S4.

Genotypes and candidate genes

All DNA was isolated using a standard CTAB extraction method [57]. DNA sequence analysis was performed using the BigDyeH Terminator Cycle Sequencing kit according to manufacturer's instructions (Applied Biosystems, Foster City, CA, USA) and resolved on an ABI3730 Capillary Sequencer at the Cornell University Life Sciences Core Laboratory Center. Twenty-one candidate genes were successfully amplified from the 27 DL subset of the association panel, using 2 or more 600 bp amplicons. The amplicon with highest nucleotide diversity was selected for sequencing in the full association panel; all DNA sequences have been submitted to GenBank as entries GF102441 through GF107318 (4,878 sequences). Sequences were assembled using Biolign 4.0.7 [58]. These genes are named in Table 1; informative polymorphisms are listed in Table S2.

F2 populations were genotyped only for the loci that were expected to segregate in each cross. Molecular markers were developed from sequence analysis of each locus and evaluated using standard PCR methods on agarose gels for indels or by fluorescently labeled primers for SNPs (Table S5). Marker data was collected and organized using Genemapper software V4.0 (Applied Biosystems, Foster City, CA, USA).

Primers were designed based on reference sequences obtained from the Maize Assembled Genomic Island (MAGI) Database [45]. Genes of interest were placed on the physical-genetic map of maize using the BLAST tool implemented by the Maize Genome Sequencing Project [42]. Gene architecture predictions were made using the FGenesH tool as implemented by Softberry [46].

Statistical tests

TASSEL 1.9.6 was used to evaluate linkage disequilibrium (LD), extract polymorphic sites, calculate narrow sense heritability, and perform General and Mixed Linear Models (GLM, MLM) with incorporation of trait data, population structure (Q) and kinship matrix (K) [59]. All other statistical analyses were done using SAS version 9.1. A t-test was used to analyze differences between NRG in the association panel under control and Al stress.

Association mapping

The MLM approach and estimation of the kinship matrix (K) has been previously described [31]. Population structure estimates (Q) have been previously described [34]. The complete results from MLM appear as Table S3. The mixed model used, for vector of phenotypes, y, is:

$$y = X\beta + Z\mu + e$$

where all fixed effects are modeled in the XB term, including genotypes and Q. Random effects are modeled in the Zm term, including the matrix of kinship coeffiecients, K, and vector of polygene background effects. e is a vector of residual effects. This model is referred to as the Q+K model. Addition of Control NRG as a fixed effect covariate in the model is referred to as the Q+K+C model.

Polymorphic sites tested, SNPs and indels, that occurred $10% were extracted from aligned sequence data. A total of 331 sites were across 21 genes were used (Table S2). Sites for the AUX1 locus were reduced to only those not in complete LD (R2 = 1) due to an excessive (73) number of sites in LD. Lines with quality scores less than 60% were discarded.

FDR

In order to account for expected false positives present due to multiple testing, a False Discovery Rate (FDR) was calculated using 1095 SNPs that occur randomly across the maize genome [60]. FDR 0:01 was calculated as:

$$FDR_{0:01} = 1 - [(X-Y)/Y]$$

Where X is the proportion of sig sites from the candidate genes # the significance value specified ($P \leq 0.01$). Y is the proportion of sig sites from the 1095 random SNPs ≤ the significance value specified. Significant sites were calculated from GLM or MLM using NRG LSmeans as the trait value.

Heritability

Marker based narrow sense heritability (h^2) was calculated in TASSEL using the kinship matrix (K) as a parent-offspring regression. Broad sense heritability (H^2) was calculated in SAS as: $H^2 \sim VG=VP$ where VG is the total genotype variance and VP is the total phenotypic variance.

Supporting Information

Table S1 Net seminal root growth data. Least Squares means were calculated for net root growth (mm 2d-1) for the association panel in the Al-stress condition (''Lmeans-Al treatment'', based on 5 replicate experiments) and control condition (''Lsmean-control'', based on 3 replicate experiments). Found at: doi:10.1371/journal.pone.0009958.s001 (0.03 MB PDF)

Table S2 Sequence polymorphisms utilized for association analysis. Polymorphic sites (SNPs and indels) were identified in each of the 21 candidate Al tolerance genes across the 282 member association panel. SNPs are coded as nucleotides (ACGT), indels are coded as numbers (e.g., 0 vs. 2), while missing data appear as N. Found at: doi:10.1371/journal.pone.0009958.s002 (0.61 MB PDF)

Table S3 Mixed Linear Model (MLM) based association analysis. MLM analysis was used to evaluate the importance for each polymorphic site in every candidate Al tolerance gene for NRG. Al-stress and control growth conditions were evaluated separately. All results are reported here. Found at: doi:10.1371/journal.pone.0009958.s003 (0.18 MB PDF)

Table S4 Validation of association mapping via linkage mapping. Association mapping results were validated using linkage mapping of F2 populations segregating for the candidate Al tolerance genes. This table reports phenotypic and genotypic information for the linkage experiments. Found at: doi:10.1371/journal.pone.0009958.s004 (0.05 MB PDF)

Table S5 PCR primers utilized for linkage mapping. Found at: doi:10.1371/journal.pone.0009958.s005 (0.02 MB PDF)

Acknowledgments

The authors would like to acknowledge the technical support provided by: Peter Bradbury, Aimee Clark, Randy Clark, Eric Craft, Jin Dong, Michael Gore,

Victoria Hodgkinson, Justin Miller, Jennifer Pleiman, Michael Rutzke, Jon Shaff, Lucas Thornblade and Zhiwu Zhang.

Author Contributions

Conceived and designed the experiments: AMK MK LVK EB OH. Performed the experiments: AMK MK. Analyzed the data: AMK MK EB OH. Wrote the paper: AMK LVK EB OH.

REFERENCES

1. Uexku¨ ll HR, Mutert E (1995) Global extent, development and economic impact of acid soils. In Date RA, Grundon NJ, Raymet GE, Probert ME, eds. Plant-Soil Interactions at Low pH: Principles and Management. Dordrecht: Kluwer Academic Publishers. pp 5–19.
2. Pandey S, Ceballos H, Magnavaca R, Bahia Filho AFC, Duque-Vargas J, et al. (1994) Genetics of Tolerance to Soil Acidity in Tropical Maize. Crop Science 34: 1511–1514.
3. Welcker C, The C, Andreau B, De Leon C, Parentoni SN, et al. (2005) Heterosis and Combining Ability for Maize Adaptation to Tropical Acid Soils: Implications for Future Breeding Strategies. Crop Science 45: 2405–2413.
4. Rao IM, Zeigler RS, Vera R, Sarkarung S (1993) Selection and Breeding for Acid-Soil Tolerance in Crops. BioScience 43: 454.
5. Vargas-Duque J, Pandey S, Granados G, Ceballos H, Knapp E (1994) Inheritance of Tolerance to Soil Acidity in Tropical Maize. Crop Science 34: 50–54. Al Tolerance Genes in Maize
6. Granados G, Pandey S, Ceballos H (1993) Response to Selection for Tolerance to Acid Soils in a Tropical Maize Population. Crop Science 33: 936–940.
7. Kochian LV (1995) Cellular Mechanisms of Aluminum Toxicity and Resistance in Plants. Annual Reviews of Plant Physiology and Plant Molecular Biology 46: 237–260.
8. Poschenrieder C, Gunse B, Corrales I, Barcelo J (2008) A glance into aluminum toxicity and resistance in plants. Science of the Total Environment 400: 356–368.
9. Delhaize E, Ryan PR (1995) Aluminum Toxicity and Tolerance in Plants. Plant Physiology 107: 315–321.
10. Ryan PR, Shaff JE, Kochian LV (1992) Aluminum Toxicity in Roots: Correlation among Ionic Currents, Ion Fluxes, and Root Elongation in Aluminum-Sensitive and Aluminum-Tolerant Wheat Cultivars. Plant Physiology 99: 1193–1200.
11. Kollmeier M, Felle HH, Horst WJ (2000) Genotypical Differences in Aluminum Resistance of Maize Are Expressed in the Distal Part of the Transition Zone. Is Reduced Basipetal Auxin Flow Involved in Inhibition of Root Elongation by Aluminum? Plant Physiology 122: 945–956.
12. Jones DL, Blancaflor EB, Kochian LV, Gilroy S (2006) Spatial coordination of aluminium uptake, production of reactive oxygen species, callose production and wall rigidification in maize roots. Plant, Cell and Environment 29: 1309–1318.
13. Kochian LV, Hoekenga OA, Pin˜eros MA (2004) How do crop plants tolerate acid soils? Mechanisms of aluminum tolerance and phosphorous efficiency. Annual Reviews of Plant Biology 55: 459–493.
14. Yamamoto Y, Kobayashi Y, Matsumoto H (2001) Lipid Peroxidation Is an Early Symptom Triggered by Aluminum, But Not the Primary Cause of Elongation Inhibition in Pea Roots. Plant Physiology 125: 199–208.

15. Jones DL, Shaff JE, Kochian LV (1995) Role of calcium and other ions in directing root hair tip growth in Limnobium stoloniferum. Planta 197: 672–680.
16. Ma JF, Ryan PR, Delhaize E (2001) Aluminum tolerance in plants and the complexing role of organic acids. Trends in Plant Science 6: 273–8.
17. Ryan PR, Delhaize E, Jones DL (2001) Function and mechanism of organic anion exudation from plant roots. Annual Reviews of Plant Physiology and Plant Molecular Biology 52: 527–560.
18. Hoekenga OA, Maron LG, Pin˜ eros MA, Canc¸ado GMA, Shaff JE, et al. (2006) AtALMT1, which encodes a malate transporter, is identified as one of several genes critical for aluminum tolerance in Arabidopsis. Proc Acad Natl Sci USA 103: 9738–9743.
19. Delhaize E, Gruber B, Ryan PR (2007) The roles of organic anion permeases in aluminium resistance and mineral nutrition. FEBS Letters 581: 2255–2262.
20. Magalhaes JV, Liu J, Guimaraes CT, Lana UG, Alves VM, et al. (2007) A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. Nature Genetics 39: 1156–61
21. Pin˜ eros MA, Shaff JE, Manslank HS, Alves VM, Kochian LV (2005) Aluminum Resistance in Maize Cannot Be Solely Explained by Root Organic Acid Exudation. A Comparative Physiological Study. Plant Physiology 137: 231–241.
22. Eticha D, Stass A, Horst WJ (2005) Cell-wall pectin and its degree of methylation in the maize root-apex: significance for genotypic differences in aluminium resistance Plant. Cell & Environment 28: 1410–1420.
23. Edmeades G, Beck D (1995) Physiology and Stress Resistant Maize Subprogram Accessed from: http://www.cimmyt.org/Research/Maize/Revisar/htm/ MFSSRB.HTM (Verified November 10, 2009).
24. Pandey S, Alberto L, Leon N, Keith D, Stephen F, et al. (2007) Breeding Maize for Tolerance to Soil Acidity. Plant Breeding Reviews 28: 59–100.
25. Garvin D, Carver B (2003) Role of the genotype in tolerance to acidity and aluminum toxicity. In Rengel Z, ed. Handbook of Soil Acidity. New York:Marcel Dekker. pp 387–406.
26. Sibov ST, Gaspar MJ, Ottoboni LMM, Arruda P, Souza AP (1999) Two genes controlling aluminum tolerance in maize: genetic and molecular mapping analyses. Genome 42: 475–482.
27. Ninamango-Ca´rdenas F, Guimaraes CT, Martins P, Parentoni SN, Carneiro NP, et al. (2003) Mapping QTLs for aluminum tolerance in maize. Euphytica 130: 223–232.
28. Mason PA (2005) Molecular and genetic investigations of aluminum tolerance in wheat and maize. Ithaca NY: Cornell University Ph.D. dissertation. 141 p.
29. Holland JB (2007) Genetic architecture of complex traits in plants. Current Opinion in Plant Biology 10: 156–161.
30. Buckler ES, Thornsberry JM (2002) Plant molecular diversity and applications to genomics. Current Opinion in Plant Biology 5: 107–11.

31. Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, et al. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nature Genetics 38: 203–8.
32. Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, et al. (2001) Dwarf8 polymorphisms associate with variation in flowering time. Nature Genetics 28: 286–9.
33. Whitt SR, Buckler ES (2003) Using natural allelic diversity to evaluate gene function. Methods in Molecular Biology 236: 123–40.
34. Flint-Garcia SA, Thuillet AC, Yu J, Pressoir G, Romero SM, et al. (2005) Maize association population: a high-resolution platform for quantitative trait locus dissection. The Plant Journal 44: 1054–64.
35. Urrea-Gomez R, Ceballos H, Pandey S, Bahia Filho AFC, Leon LA (1996) A Greenhouse Screening Technique for Acid Soil Tolerance in Maize. The Agronomy Journal 88: 806–812.
36. Magnavaca R, Gardner C, Clark R (1987) Evaluation of inbred maize lines for aluminum tolerance in nutrient solution. In Gabelman H, Loughman BC, eds. Genetic Aspects of Plant Mineral Nutrition. Dordrecht: Martinus Nijhoff. pp 255–265.
37. Pin˜ eros MA, Magalhaes JV, Alves VM, Kochian LV (2002) The physiology and biophysics of an aluminum tolerance mechanism based on root citrate exudation in maize. Plant Physiology 129: 1194–206.
38. Maron LG, Kirst M, Mao C, Milner MJ, Menossi M, et al. (2008) Transcriptional profiling of aluminum toxicity and tolerance responses in maize roots. New Phytologist 179: 116–128.
39. Sasaki T, Yamamoto Y, Ezaki B, Katsuhara M, Ahn SJ, et al. (2004) A wheat gene encoding an aluminum-activated malate transporter. The Plant Journal 37: 645–653.
40. Magalhaes JV, Garvin DF, Wang Y, Sorrells ME, Klein PE, et al. (2004) Comparative mapping of a major aluminum tolerance gene in sorghum and other species in the poaceae. Genetics 167: 1905–14.
41. Fontecha G, Silva-Navas J, Benito C, Mestres MA, Espino FJ, et al. (2007) Candidate gene identification of an aluminum-activated organic acid transporter gene at the Alt4 locus for aluminum tolerance in rye (Secale cereale L.). Theor Appl Genet 114: 249–60.
42. Liu J, Magalhaes JV, Shaff J, Kochian LV (2009) Aluminum-activated citrate and malate transporters from the MATE and ALMT families function independently to confer Arabidopsis aluminum tolerance. Plant J 57: 389–99.
43. Wang J, Raman H, Zhou M, Ryan PR, Delhaize E, et al. (2007) High-resolution mapping of the Alp locus and identification of a candidate gene HvMATE controlling aluminum tolerance in barley (Hordeum vulgare L.). Theor Appl Genet 115: 265–76.
44. The Maize Genome Sequencing Project accessed from http://www.maizesequence. org (Verified November 10, 2009).
45. Fu Y, Emerich SJ, Guo L, Wen TJ, Ashlock DA, et al. (2005) Quality assessment of maize assembled genomic islands (MAGIs) and experimental

verification of predicted novel genes. Proc Acad Natl Sci USA 102: 12282–12287.
46. Softberry, Inc. Prediction of protein secondary structure using Markov chains. Accessed from http://linux1.softberry.com/berry.phtml?topic = pps&group = programs&subgroup =propt (Verified November 10, 2009).
47. Ravanel S, Gakiare B, Job D, Douce R (1998) The specific features of methionine biosynthesis and metabolism in plants. Proc Acad Natl Sci USA 95: 7805–7812.
48. Detarsio E, Maurino V, Alvarez C, Mu¨ ller C, Andreo C, et al. (2008) Maize cytosolic NADP-malic enzyme (ZmCytNADP-ME): a phylogenetically distant isoform specifically expressed in embryo and emerging roots. Plant Molecular Biology 68: 355–367.
49. Li CH, Yu N, Jiang Sm, Shangguan XX, Wang JL, et al. (2008) Downregulation of S-adenosyl-L-homocysteine hydrolase reveals a role of cytokinin in promoting transmethylation reactions. Planta 228: 125–36.
50. Weretilnyk EA, Alexander KJ, Drebenstedt M, Snider JD, Summer PS, et al. (2001) Maintaining Methylation Activities during Salt Stress: The Involvement of Adenosine Kinase. Plant Physiology 125: 856–865.
51. Yang JL, Li YY, Zhang YJ, Zhang SS, Wu YR, et al. (2008) Cell Wall Polysaccharides Are Specifically Involved in the Exclusion of Aluminum from the Rice Root Apex. Plant Physiology 146: 602–611.
52. Harjes CE, Rocheford TR, Bai L, Brutnell TP, Kandianis CB, et al. (2008) Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. Science 319: 330–3.
53. Ligaba A, Katsuhara M, Ryan PR, Shibasaka M, Matsumoto H (2006) The BnALMT1 and BnALMT2 Genes from Rape Encode Aluminum-Activated Malate Transporters That Enhance the Aluminum Resistance of Plant Cells. Plant Physiology 142: 1294–1303.
54. Pin˜ eros MA, Canc¸ado GMA, Maron LG, Lyi SM, Menossi M, et al. (2008) Not all ALMT1-type transporters mediate aluminum-activated organic acid responses: the case of ZmALMT1 - an anion-selective transporter. The Plant Journal 53: 352–67.
55. Pin˜ eros MA, Canc¸ado GMA, Kochian LV (2008) Novel Properties of the Wheat Aluminum Tolerance Organic Acid Transporter (TaALMT1) Revealed by Electrophysiological Characterization in Xenopus Oocytes: Functional and Structural Implications. Plant Physiology 147: 2131–2146.
56. Kovermann P, Meyer S, Ho¨rtensteiner, Picco C, Scholz-Starke J, et al. (2007) The Arabidopsis vacuolar malate channel is a member of the ALMT family. The Plant Journal 52: 1169–1180.
57. Saghai-Maroof MA, Soliman K, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA 81: 8014–8018.
58. Hall T, Biolign 4.0.7 (computer program). Accessed from http://www2. maizegenetics.net/bioinformatics (Verified November 10, 2009).

59. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, et al. (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23: 2633–5.
60. Yamasaki M, Tenaillon MI, Vroh Bi I, Schroeder SG, Sanchez-Villeda H, et al. (2005) A large-scale screen for artificial selection in maize identifies candidate agronomic

CHAPTER 3

NESTED ASSOCIATION MAPPING FOR ALUMINUM TOLERANCE

Introduction

Two main approaches for identifying markers linked to traits of importance in breeding are linkage mapping and association mapping. These approaches vary in type and number of individuals and amount of marker data needed. There are advantages and disadvantages to each approach, but they can also complement each other when used together.

Linkage mapping has historically been used in a majority of maize quantitative trait studies and is usually based on a single experimental cross of two inbred lines. It is powerful tool for identifying chromosomal regions with major effects on the trait between those parental lines. Populations of F_2 or backcrosses families can be used, but recombinant inbred lines (RIL) are most often used, due to their low heterozygosity, greater recombination and fixed genetics. The use of inbred lines with fixed genomes provides a powerful resource for replication in different environments in order to separate the genetic and environmental effects on a phenotype. Linkage mapping identifies quantitative trait loci (QTL) and linked markers segregating between two parental haplotypes with limited recombination. Limited recombination gives rise to large chromosomal regions in high linkage disequilibrium (LD), reducing the number of markers needed, but also making it difficult to narrow down a region for identifying specific genes. Alleles in linkage mapping studies are at high frequency, but linkage mapping can only identify QTL segregating in a specific cross. Often this represents only a small fraction of the QTL and phenotypic variation in a species. Many different mapping populations are needed to represent the allelic diversity of the genes contributing to the trait of interest [1].

Association mapping, or population mapping, uses historical recombination events that have occurred during the evolution of a diverse population. This greatly increases both the resolution and the number of alleles that can be identified, compared to just two segregating alleles per locus in a biparental population, but association mapping studies also requires a much greater amount of genetic data. Since the LD in a diverse association population is much lower than in a biparental population, far more markers are required to ensure that at least one is linked to a functional variant. There are different rates of LD decay in different species. Outcrossers, like maize, have LD that breaks down more rapidly than an inbreeder or plants more likely to self pollinate (e.g. *Arabidopsis*, rice). In maize, LD (r^2) generally decays to 0.2 within about 1kb in landraces, 2kb in inbred lines, and 500 kb in elite inbred lines, although it varies across loci depending on selection pressure [1]. Association mapping takes advantage of random mating over generations, but in most cases nonrandom mating has also occurred, creating population structure. This population must be accounted for in phenotype-genotype association models to minimize false associations. False associations can also

arise though indirect associations, such as links between causal and non causal sites, multiple causal sites or epistasis [2]. False discovery rates (FDR) can be calculated from random markers as a percentage of false associations that can be expected with particular phenotype.

Early association mapping studies were limited by genotyping costs and were most useful for testing candidate genes or fine mapping a QTL. This approach has been used in maize, lettuce, barley, wheat, oat, soybean and triticale [3]. It has been successful in identified genes controlling or contributing to quantitative traits in maize, such as vitamin A, beta-carotene, kernel composition, starch production, flowering time, and Al tolerance. [1, 4-6]. Association mapping typically identifies high frequency alleles of small effect (explaining <5% of variation), which tend to control a majority of maize traits including flowering time, kernel composition, southern leaf blight, oil concentration, and leaf architecture [1]. However, association mapping may miss larger effect alleles, at relatively low frequency, such as *ZmMATE1*.

Lowered costs of genetic data and high throughput sequencing have made genome wide association studies (GWAS) more feasible. GWAS essentially covers the entire genome with a dense amount of genetic markers. It is estimated that 10-15 million SNP markers may be necessary to conduct a thorough GWAS in a diverse maize panel [1, 7]. However, even this amount of coverage does not guarantee coverage of other genetic polymorphisms often associated with a phenotype, such as indels, transposons, copy number variations, or epigenetic effects. GWAS can be used to compliment QTL studies, but associated polymorphisms may not be explaining a sufficiently large fraction of phenotypic variance to withstand multiple test corrections. Cook et al found no significant sites for maize kernel composition when using GWAS in an association panel, but a candidate gene approach did find significance polymorphism [6].

The Maize Nested Association Mapping (NAM) recombinant inbred line population is a collection of 5000 recombinant inbred lines developed by the NSF-funded Molecular and Functional Diversity of Maize project and the Maize Diversity Group [8]. This population combines the advantages of both linkage and association mapping and was created as a publically available resource to improve mapping resolution without requiring the dense markers of a GWAS study. It links 25 RIL populations by a common parent, B73, which has a published physical map and full genomic sequence. Each population has ~200 RILs for a total of ~5000 lines, which capture roughly 80% of the diversity found in maize. Extensive sequencing on the 25 parental lines combined with low density sequencing on the RILS allows for projection of parental genotypic information onto the entire population[9]. This NAM population was used for a preliminary study of aluminum (Al) tolerance in maize.

Materials and Methods

This experiment was conducted prior to completion and release of the entire NAM population, therefore not all ~5000 RILs were not used. A total of 4130 RILs including 254 RILs from the IBM population (B73x Mo17) were used. An average of

150 lines per population were used, with five populations having fewer than 100 lines (Figure 2). Plants were phenotyped as previously described for the association study, with a two deviations (Chapter 2) [5, 10]. First, only two seeds per line were germinated and only one plant per RIL was used in the experiment. Second, a higher Al^{3+} activity of 38μM was used instead of 27μM. This higher level of Al^{3+} is frequently used in other mapping studies for maize Al tolerance [11-13]. RILs within populations were randomized, but populations were kept together within blocks of 3- 5 hydroponic tubs. Each tub contained 2 lines each of B73, the non-B73 parental inbred, and Cateto100-6 (an Al tolerant line frequently used in mapping studies [11, 14]). These lines were included as checks to determine tub effects. The experiment was done over several weeks during the summer of 2006.

Three phenotypic values were used in the analysis, Net Root Growth (NRG), Relative Root Growth (RRG) and Best Linear Unbiased Prediction (BLUP) estimates of NRG. RRG was calculated as (FRG-IRG)/IRG. All negative NRG and RRG values were represented as missing data. BLUPs were estimated for NRG using a mixed linear model (MLM) procedure in SAS, using line, population and tub(population) as fixed/random effects (Figure 2). LSMeans of the population were added to these BLUP estimates. Joint linkage QTL mapping was performed by stepwise regression for the three phenotypes (NRG, RRG, and BLUP) using GLM Select in SAS as previously described [15]. Narrow sense heritability (h^2) was calculated for NRG and RRG as the regression coefficient of the population means on the mid-parent values (Table 1). Although the IBM population was phenotyped as part of this experiment, it was not included in the analysis (pop 17), because at the time IBM RILs had not yet been incorporated into to the NAM population.

Results

Three trait values (NRG, BLUP-NRG, and RRG) were used for analysis and varied in their distributions among and between populations. NRG did not follow a normal and was severely skewed toward low root growth, whereas RRG and BLUP-NRG estimates followed a more normal distribution (Figure 1). This pattern was constant both across and within populations (Figure 2). There were several roots with negative NRG (4%) and no (6%) NRG. Negative NRG values were most likely a result of broken roots and were represented as missing data. Narrow sense heritability (h^2) was greater for RRG (76%) than NRG (48%) (Table 1).

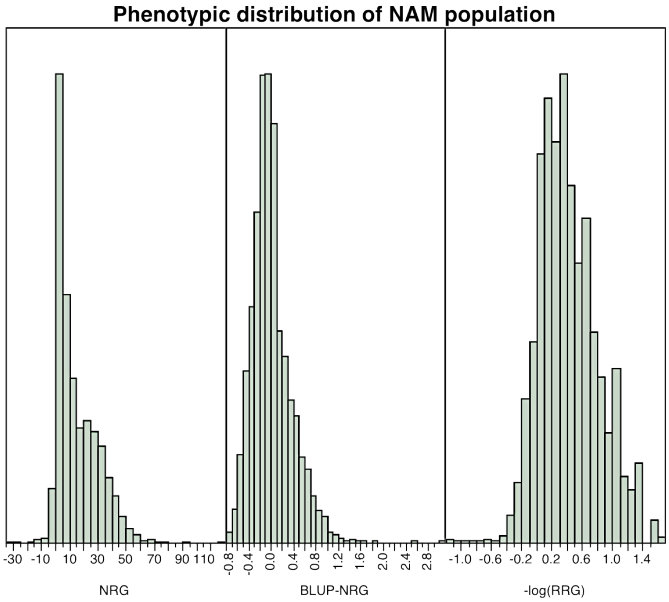

Figure 1. Histograms showing the distribution of traits, NRG, RRG, and BLUP_NRG, in the entire NAM population.

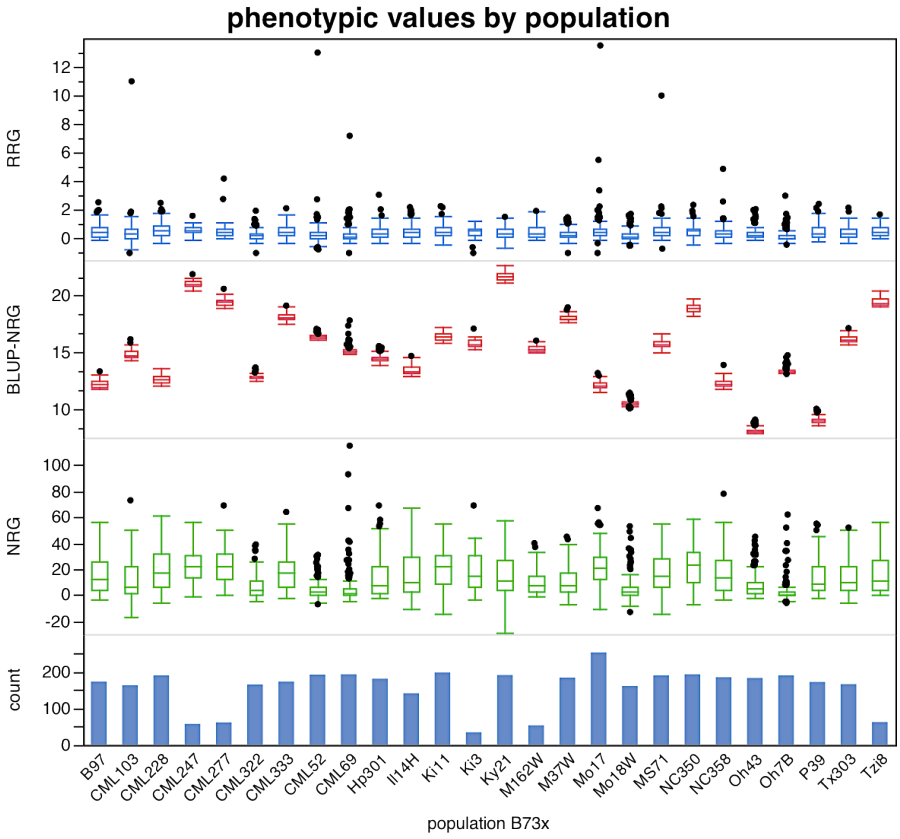

Figure 2. Phenotypic distributions by population for NRG, BLUP-NRG and RRG and number of lines used in each.

GLMSelect was used to identify 21 unique markers covering 15 QTL BINs across 7 chromosomes for all traits combined, as described previously [15]. NRG, BLUP-NRG, and RRG, and had 8, 5 and 9 markers significant at p<0.01, respectively, explaining 26%, 7% and 16% of the total phenotypic variance (R^2), respectively (Table 1 and Figure 3).

Table 1. Number of markers identified in the model (p <0.01) and r^2 for each phenotype, not including negative root growth

Trait	# of markers in model	R^2	h^2 (heritability)
NRG	8	26%	48%
BLUP-NRG	5	7%	
RRG	9 (3 removed)	16%	73%

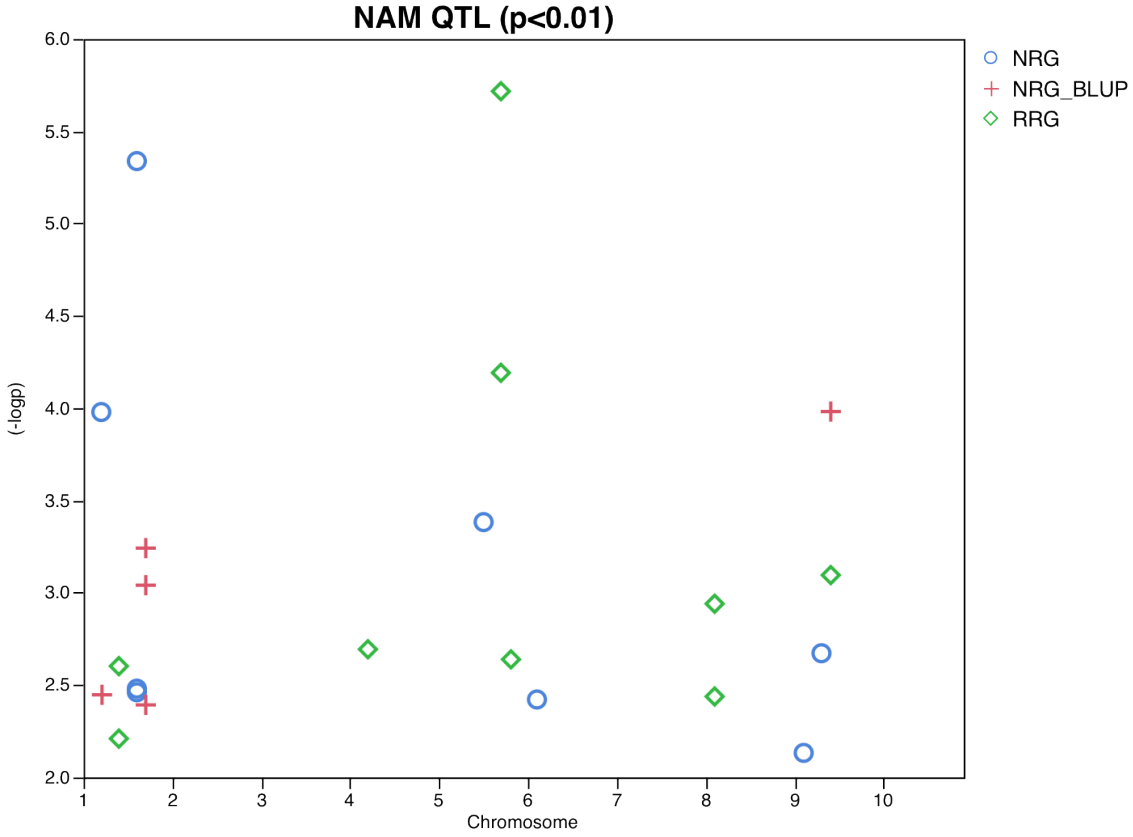

Figure 3. Significant markers (p<0.01) for NRG, RRG, and BLUP estimates (shown as –log p values) by chromosome (negative root growth values not included)

Eight unique markers were identified on chromosome 1 across 4 QTL BINs (1.02, 1.04, 1.06 and 1.07). Marker 88 in QTL 1.06 for NRG had the second highest significance value (p= 4.58E-06). Only one marker was identified on chromosome 4 (440) in BIN 4.02 for RRG. Four markers were identified on chromosome 5 across 3 QTL BINs (5.05, 5.07 and 5.08). The marker of highest significance was located at marker 677 (p= 1.91E-06) under QTL BIN 5.07 for RRG. One marker (700) on chromosome 6 for NRG falls within a major QTL BIN (6.01) found in previous studies. Two markers were identified on chromosome 8 (8.01) for RRG and three makers were identified on chromosome 9 (9.01, 9.03, 9.04) for NRG, BLUP-NRG and RRG, respectively. Marker 999 in BIN 9.04 was the only marker identified for more than one

phenotypic value (RRG and BLUP-NRG). Two markers were also found on chromosome 10 (10.04, and 10.07).

When negative root growth values are included in the model it has a significant effect on the results (Table2, Figure 4). 28 markers ($p<0.01$) are identified in this model, but only 8 overlap with those identified previously (negative root growth removed, Figure 3). There are still a large number of markers (9) on Chromosome 1 across all traits. However, several markers (4) show up on chromosome 6 (BIN 6.01) under RRG and one under BLUP-NRG. The most significant markers 695 ($p= 7.59E-54$) is identified for RRG.

Table 2. Number of marker effects in the model ($p <0.01$) and r^2 for each phenotype, including negative root growth

Trait	# of markers in model	r^2
NRG	10	29%
BLUP-NRG	9 (1 removed)	11%
RRG	9 (3 removed)	26%

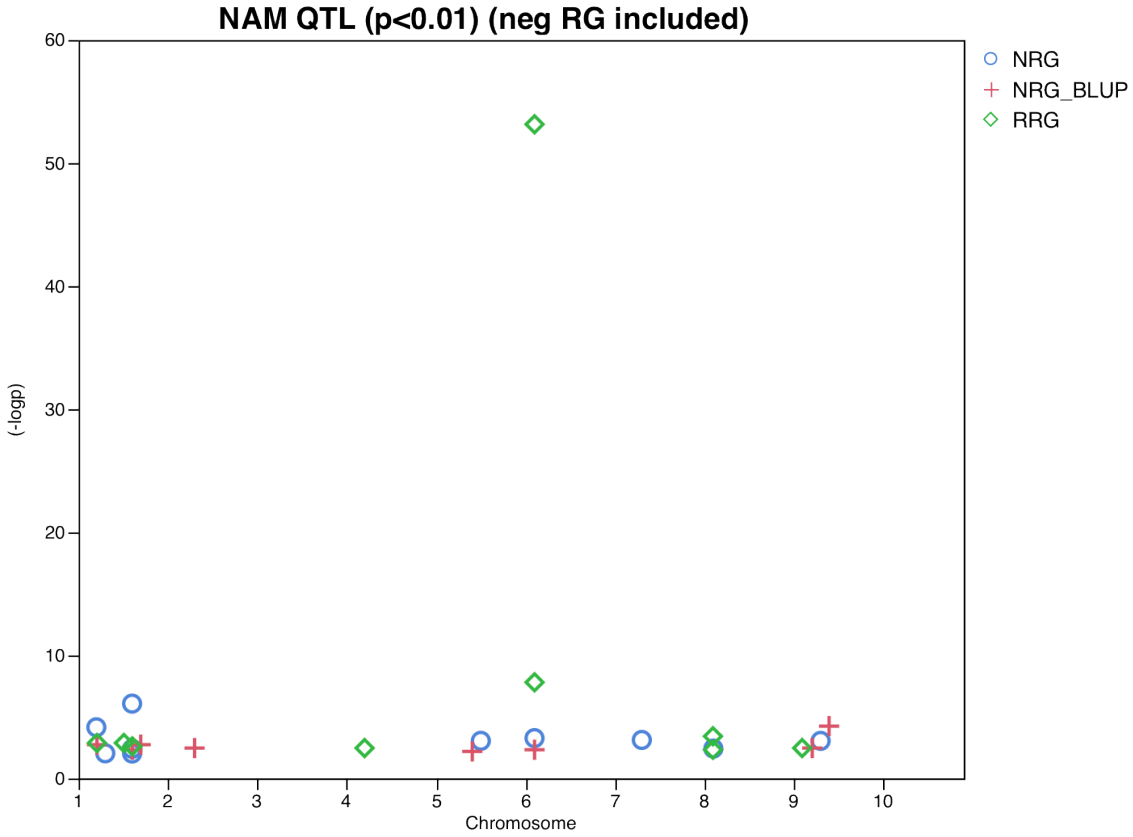

Figure 4. Significant markers ($p<0.01$) for NRG, RRG, and BLUP estimates (shown as –log p values) by Chromosome (negative root growth values included)

Discussion

Although heritability estimates are fairly high, these results should be taken as very preliminary for several reasons. First n=1 was a problem because there was not sufficient replication of the NAM RILs and there were several issues with phenotyping. Second, the choice of 38μM Al^{3+} instead of 27μM Al^{3+} was poor because a majority of lines were highly sensitive to the higher concentration, which skewed the phenotypic distribution toward low or no root growth. Third, statistical analysis showed a very large tub effect contributing to root growth, based on the replicated checks within each tub. A fungus was present in many tubs, which may have further stressed the roots and suggests a contamination of the tubs, tools, or solution. The roots were measured by hand, which was done by removing the small plastic container housing the maize seedling and holding a ruler along the seminal root. Breakage of the root during these measurements likely accounts for the large negative root growth measurement observed. In order to get an accurate measure of Al toxicity and root growth for the NAM population, several more replications would be needed to separate out these environmental and experimental factors.

Only one marker (700) is located within any QTL previously identified. This QTL is indentified as by Sibov et al as *Alm1* (6.01) (Chapter 1, Table 2)[16]. Also the expression studies of genes upregulated by Al also linked several genes to these QTL [17]. Most of the QTL identified are novel, which is expected when looking across several different linkage populations not previously used in Al tolerance studies.

A few QTL overlap with candidate genes used in the association and linkage study (Chapter 2)[5]. Two genes that correlated with Al tolerance in this study fall within two QTL identified in the NAM population, *ZmASL* and *ALMT2 (*1.02 and 10.04, respectively*),* but both are several Mbps away from the markers identified with closer markers not being significant. Two candidate genes, *fumerase* and *germin2* (*oxalate oxidase*), that were chosen for the association study based on their increased expression in root tips under Al stress, but were not significant in the association studies in Chapter 2, fall within QTL 1.04 and 10.03, respectively.

Including negative root values creates a highly significant marker. The parameter estimates of this highly significant marker (695) in QTL 6.01, show population 8 (CML52) as the main contributor with a p value of $1.08E^{-102}$. Population 22 (Oh43) is the only other population significant at $p<0.01$ for this marker. Population 8 also contributes significantly to a nearby marker 693 ($p= 5.90E^{-104}$). This marker has a p value of $1.49E^{-08}$ in the stepwise model. In population 8, 9% of the lines have negative values, and there is one major outlier, which is not caused by negative root growth and therefore is in both models. It is possible this is a locus related to brittle roots or lignin deposition caused by the toxic effects of Al. Lignin biosynthesis enzymes were differentially expressed between tolerant and sensitive lines in expression studies.

Conclusions

Al tolerance in maize is a complex quantitative trait controlled by many genes or loci. A combination of physiological, genetic, and comparative genomic approaches has been crucial in the discovery Al tolerance genes in several species including maize. Organic acid exudation as an Al tolerance mechanism has been observed across many members of the Poaceae family, as well as, other species. However, two highly Al tolerance grasses, rice and a forage grass (*Bracchiaria decumbens*), show no evidence for Al activated organic acid exudation as a mechanism for tolerating Al. Many Al tolerance genes that have been identified are membrane transporters involved in Al activated organic acid efflux, including *ZmMATE1* in maize, a citrate transporter. However, in maize there is evidence that Al activated citrate release is not the only tolerance mechanism.

Citrate release plays an important role in Al tolerance in maize, but unlike wheat, where there is a strong correlation between tolerant varieties and malate release, the correlation between citrate exudation and Al tolerance lacks strong evidence. The correlation between citrate exudation from roots and increased tolerance in maize was initially studied in only a single Al tolerance variety and one or two Al sensitive varieties. Pineros et al [18] tested this correlation in a broader panel of six diverse maize genotypes varying in tolerance levels (Cateto, Pioneer 3355, Mo17, B73, L53, 11x723). All genotypes except B73 (moderately Al-sensitive) show an Al activated increase in citrate exudation and Mo17 (moderately Al-sensitive) exhibited the highest rate. Although all the tolerant varieties tested so far do exhibit a significant amount of Al induced citrate release, there was no correlation when looking across diverse genotypes [18]. This suggests maize has mechanisms in addition to citrate release to tolerate Al or that citrate release may work in conjunction with other mechanisms or genes.

ZmMATE1 was the first gene positively identified as contributing toward Al tolerance in maize. The discovery of the functional variant in the *ZmMATE1* gene, a copy number variation (CNV), is a significant step toward incorporating tolerance into elite germplasm. The *ZmMATE1* CNV allele contributing to tolerance is a relatively rare allele and therefore would not have been detected in an association study. If rare alleles are the main contributors to Al tolerance, as they are in other quantitative traits, more comprehensive studies using the NAM and similar populations could be the most efficient way to identify genes. NAM has more power than an association population for detecting small effect loci and rare alleles. It can also give us an idea of how genes are interacting in different populations or backgrounds [7].

Discovery of more Al tolerance genes could help to increase levels of tolerance, possibly in conjunction with *ZmMATE1* or by providing alternative mechanisms for those backgrounds where citrate release does not contribute to an increase in tolerance. The lines with the *ZmMATE1* allele with the CNV/tandem triplication contributing to Al tolerance were all endemic to regions with high acid soils. Maize landraces and inbred lines adapted to these high acid soils could be a valuable resource for gene mining due to directional selection for Al tolerance. The heterogeneity of the landraces can be both a strength and weakness. These landraces are a source of tremendous genetic diversity that has not been subjected to the recent bottleneck of elite inbred line development.

However, these landraces perform relatively poorly in the field when compared to these elite lines. This heterogeneity could provide insight into the relationship between alleles contributing to Al tolerance, but would also make dissection of specific loci more difficult.

The candidate genes identified in Chapter 2 (based on association and linkage studies) as having a positive association and linkage (*ALMT2, ZmASL, SAHH, and ME*) are still excellent candidates as potential Al tolerance genes. Also those genes with a positive association to root growth but not found to be significant in the linkage populations (PME, ISL and ME), could still be candidates since expression studies showed an up-regulation of these genes in response to Al. The associated sites could be in LD with a polymorphisms regulating expression, which the linkage populations did not segregate for.

ZmALMT2 is the only one of these candidate genes that has been further studied and functionally characterized, however the significant alleles from the association and linkage study were not included so the effect of these alleles are still unknown. *ZmALMT2* is characterized as a plasma membrane transporter that mediates a bidirectional, highly selective transport of both organic and inorganic ions, and facilitates a large constitutive root malate efflux, regulated in part by the cytoplasmic and extracellular environment. Although it is the first ALMT in maize to be identified as a transporter of malate outside the cell, this study suggest *ZmALMT2* is not involved in Al tolerance due to several characteristics that are different from those seen in other ALMTs involved in Al tolerance. First, expression is greater in the mature root region than the root tip. Second, the transporter is not activated by Al and is, in fact, suppressed by it. Third, expression is significantly higher in the sensitive B73 line than the tolerant Cateto line used in this study. In previous physiological studies, B73 was found to have no Al activated citrate or malate efflux, so Al activation of an OA release mechanism should not be expected. This suggests B73 does not use an Al activated malate or citrate release mechanism, but an alternative mechanisms to tolerate Al at a moderate level [18].

Association studies also show that B73 carried the superior dominant *ZmALMT2* allele, increasing root growth by 20% in the association population and 15% in the linkage population, when compared to the alternative significant allele (NC350 in the linkage population) [5]. The significant non-B73 (NC350) allele was not included in this study and it is not known what allele Cateto has, as it was not part of the association panel, and is not reported in the study. It is possible Cateto has the alternative allele, but its high level of tolerance from the rare *ZmMATE1* triplication could mask the effect of the *ZmALMT2* allele [19]. *ZmALMT2* is believed to play several roles in plant mineral nutrition and have a questionable role in Al tolerance [20]. However, further studies should be done to fully understand the role of the allele from the association and linkage studies before ruling this gene out. This is the first ALMT gene characterized in maize that shows a high efflux of malate, which is likely to contribute to protecting the root from Al damage or has other functions to protect the cell besides chelating Al in the rhizosphere, such as controlling malate concentrations in the cell or chelating Al inside the cell. Expression of *ZmALMT2* in transgenic Al sensitive Arabidopsis plants, that are

Al hypersensitive as both *AtALMT1* and *AtMATE* have been knocked out, showed a large efflux of malate that was also enhanced by Al, partially restoring Al tolerance. This Al enhanced malate transport was not seen when expressed in *Xenopus* oocytes, suggesting additional post-translational modifications in plants may modulate the activity of *ZmALMT2 [20]*.

ZmASL, is a gene in the MATE family with a strong correlation with Al tolerance in the association and linkage populations (Chapter 2) [5]. It is located within Bin 1.02, where significant markers were identified for both RRG and BLUP-NRG in the NAM population. However, this genes is about ~3.5Mbp away from the marker identified with closer markers having no significance. The functional characteristics of this gene are currently unknown, but it has high similarity to *HvMATE* in Barley. This gene is highly polymorphic in the association population, with many amino acid substitution and indels. 43 amino acid substitutions and well as many polymorphisms were classified as rare alleles and were not tested in the association analyses due to lack of power to test for low frequency alleles. This gene is still a good candidate for further molecular and physiological studies, as citrate release contributes to Al tolerance in many tolerant genotypes and rare alleles, such as *ZmMATE1,* are known to contribute to Al tolerance.

S-adenosyl-L-homocysteinase (*SAHH*) is an enzyme that removes a feedback inhibitor for SAM methylation, which is involved in DNA/RNA modification, nucleic acid metabolism and synthesis of cell wall constituents. SAHH was chosen as a candidate gene because it was highly expressed in root tips under Al stress. It has also been associated with a salt-stress response in spinach and sugar beets and viral resistance in Arabidopsis[21]. *SAHH* is a highly conserved gene among eukaryotes, and there were very few polymorphisms within the entire ORF. Three significant polymorphisms (SNPs) were found in the region sequenced in the association panel, which included the first exon. One of these SNPs was triallelic and caused the only amino acid substitution observed throughout the entire gene sequence in 27 diverse lines. This triallelic SNP leads to either a synonymous or conservative amino acid substitution. The contribution of SAHH to Al tolerance could come through any of several mechanisms due to the broad range of processes. For example, *SAHH* may contribute to Al tolerance through mechanisms involving cell wall modification or pectin methylation.

NADP-Malic Enzyme (ME) is a cytosolic enzyme that catalyzes the conversion of malate to pyruvate. It has been linked to plant defense responses, such as hypoxia and drought, lignin biosynthesis and control of cytosolic pH by balancing synthesis and degradation of malate [22]. This enzyme was chosen as a candidate gene due to a high up regulation by Al in expression studies [23]. It also falls within a QTL (6.05) identified in two studies [12, 24].

Many studies have strongly suggested that an important site of Al toxicity is the cell wall of the root tip. Studies differ in the correlation between Al tolerance and reduced Al accumulation in the root. In studies comparing one or two tolerant and sensitive lines, the tolerant lines exhibit significantly less Al accumulation in their root tips [23, 25]. Other studies using larger bi-parental or association populations show only

a partial correlation between Al exclusion and tolerance [13](Hoekenga, unpublished). This suggests mechanisms that keep Al from binding to the root contribute to Al tolerance, but other mechanisms to tolerate Al in the roots also exist. Organic acids or other ligands can chelate Al both in the rhizosphere and symplasm. Other possible mechanisms could involve root-mediated changes in rhizosphere pH, root exudation of Al-chelating phenolic compounds, reduction in negatively charged binding sites in either the cell wall and/or the plasma membrane, changes in methylation of pectin residues in the cell wall, or internal tolerance mechanisms, such as chelation by organic ligands. Many genes discovered via expression or association analysis may not be specific to Al exposure, but are expressed as general abiotic stress responses, and may still contribute to a plants ability to decrease toxic effects or grow in a toxic environment and should not be overlooked.

Future of Acid soils

For maize and other crops grown on acidic and Al toxic soils, amendment of soil with lime and fertilizers, in combination with Al tolerant genotypes can greatly increase yields. It is usually more economical to develop Al tolerant varieties than to correct soil deficiencies or toxicities, but the use of Al tolerant germplasm without neutralizing or sustaining soil fertility can lead to increased acidity due to the continued removal of minerals and organic matter. Sustainable management practices and inputs must be used to ensure fertility of these soils. A low input system has never been shown to work over a long-term for crops such as maize, which require higher inputs and are not as adaptable to these soils as native plants [26].

Acid soil regions in Brazil have been successfully developed for sugarcane and soybean cropping. The use of higher yielding cultivars and sustainable management practices has allowed plantation crops to thrive on acid soils for decades. Soil fertility has been replenished on degraded and deforested land that was initially cleared for crops and subsequently abandoned due to infertility. Soil fertility replenishment was achieved though a combined use of fast growing legume cover crops, lime and phosphate applications, and soil conservation. [26]

Restoring fertility to these areas is costly and time consuming, but the forests and wetlands of the tropics normally cleared for increased cropland are ecologically and biologically very important. These tropical forests are not only important for ecosystem balance and hydrological cycling, but they contain a tremendous amount of biodiversity [26]. Although modern technology has made it possible to clear and crop this land for increased food production, it may be more important to conserve this land and focus on restoration of land already deforested, abandoned and unproductive. Von Uexkull outlines a three step approach for improving these lands. It starts by establishing a leguminous cover crop and application of lime, followed by deepening and enriching the rooting zone, then efficiently balancing nutrition and management practices to establish an ecologically and economically sustainable system [26]. While this approach would contribute greatly to improving many unproductive lands, it would in itself be unproductive and expensive for the years it takes fertility to be restored. For this reason,

it will be possible only within a system that places economic value on the ecological and biological assets of undeveloped tropical wetlands and forests.

Al toxicity is a major agronomic concern, second only to drought. The use of Al tolerant varieties can help increase production of those acidic lands already cleared, reducing the need to expand into existing forests or wetlands. The most efficient was to develop Al tolerant varieties is to identification genes and mechanisms which plants use to tolerate Al. There have been many advances in tools and resources available to discover genes, not only in genetics but also in phenotyping methods. Digital imaging of plant roots provides more accurate data about the entire root system, while minimizing handling. Hydroponics provides an efficient way to measure tolerance, but does not fully capture how a plant performs in a field environment. Field studies continue to be a challenge, due to environmental effects confounding genetic effects (GxE). Development of an efficient field trial system would help to move these varieties or loci in a direction closer to use in breeding programs and close the gap between geneticists and breeders. The knowledge of Al tolerance in maize has grown considerably over the years, but there is still much to be discovered about this complex trait.

REFERENCES

1. Myles, S., et al., *Association Mapping: Critical Considerations Shift from genotyping to Experimental Design.* The Plant Cell, 2009. **21**: p. 2184-2202.
2. Larsson, S., A. Lipka, and E.S. Buckler, *Lessons from Dwarf8 on the Strengths and Weaknesses of Structured Associaion Mapping.* PLos Genetics, 2013. **9**(2).
3. Niedziela, A., et al., *Aluminum tolerance association mapping in triticale.* BMC Genomics, 2012. **13**(67).
4. Harjes, C.E., et al., *Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification.* Science, 2008. **319**(5861): p. 330-383.
5. Krill, A., et al., *Association and Linkage Analysis of Aluminum Tolerance Genes in Maize.* Plos One, 2010. **5**(4).
6. Cook, J., et al., *Genetic Architecture of Maize Kernel Compostition in the Nested Association Mapping and Inbred Association Panels.* Plant Physiology, 2012. **158**: p. 824-834.
7. Wallace, J., S. Larsson, and E.S. Buckler, *Entering the second century of maize quantitative genetics.* Heredity, 2013: p. 1-9.
8. http://www.panzea.org/, M.d.g. *Genetic Architecture of Maize and Teosinte_dev site*. 2011.
9. Michael D. McMullen, et al., *Genetic Properties of the Maize Nested Association Mapping Population.* Science, 2009. **325**.
10. Magnavaca, R., C. Gardner, and R. Clark, *Evaluation of inbred maize lines for aluminum tolerance in nutrient solution.* Genetic Aspects of Plant Mineral Nutrition, 1987. **27**: p. 225-265.
11. Maron, L.G., et al., *Two functionally distinct members of the MATE (multi-drug and toxic compound extrusion) family of transporters potentially underlie two major aluminum tolerance QTLs in maize.* The Plant Journal, 2009. **61**(5): p. 728-740.
12. Ninamango-Cárdenas, F.E., et al., *Mapping QTLs for aluminum tolerance in maize.* Euphytica 2003. **130**(2): p. 223-232.
13. Mason, P., *Molecular and genetic investigations of aluminum tolerance in wheat and maize.* , 2005, Cornell University: Ithaca NY:. p. 141 p.
14. Guimarães, C.T., et al., *QTL and selection Mapping for Al tolerane in Tropical Maize*, 2009.
15. Buckler, E.S., J.B. Holland, and M. McMullen, *The Genetci Architecture of Maize Flowering TIme.* Science, 2009. **325**.
16. Sibov, S.T., et al., *Two genes control aluminum tolerance in maize: Genetic and molecular mapping analyses.* Genome, 1999. **42**: p. 475-482.
17. Mattielloa, L., F. Rodrigues da Silvab, and M. Menossi, *Linking microarray data to QTLs highlights new genes related to Al tolerance in maize.* Plant Science, 2012. **191-192**: p. 8-15.
18. Piñeros, M., et al., *Aluminum Resistance in Maize Cannot Be Solely Explained by Root Organic Acid Exudation. A Comparative Physiological Study.* Plant Pysiology, 2005. **137**(1): p. 231-241.
19. Maron, L., et al., *Aluminum tolerance in maize is associated with higher MATE1 gene copy number.* PNAS, 2013. **110**(13): p. 5241-5246.

20. Ligaba, A., et al., *Maize ZmALMT2 is a root anion transporter that mediates constitutive root malate efflux.* Plant, Cell & Environment, 2012. **35**(7).
21. Weretilnyk, E.A., et al., *Maintaining Methylation Activities during Salt Stress. The Involvement of Adenosine Kinase.* Plant Physiology, 2001. **125**(2): p. 856-865.
22. Schaaf J, Walter MH, and H. D, *Primary metabolism in plant defense (regulation of a bean malic enzyme gene promoter in transgenic tobacco by developmental and environmental cues).* Plant Physiology, 1995. **108**: p. 949-960.
23. Maron, L.G., et al., *Transcriptional profiling of aluminum toxicity and tolerance responses in maize roots.* New Phytologist, 2008. **279**(1): p. 116-128.
24. Conceição, L.D.H.C.S., C. Tessele, and J.F. Barbosa Neto, *Diallele analysis and mapping of aluminum tolerance in corn inbred lines.* Maydica, 2009. **54**: p. 55-61.
25. Mattielloa, L., et al., *Transcriptional profile of maize roots under acid soil growth.* BMC Plant Biology 2010. **10**(196).
26. von Uexküll, H.R. and E. Mutert, *Global extent, development and economic impact of acid soils.* Plant and Soil, 1995. **171**: p. 1-15.